AD-A010 210

CONTENT ANALYSIS OF SURVEY FEEDBACK MEETINGS: AN EVALUATION TOOL

Patricia A. Pecorella

Michigan University

Prepared for:

Office of Naval Research

May 1975

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. 3. RECIPIENT'S CATALOG NUMBER: AD-A010 210
4. TITLE (and Subtitle): Content Analysis of Survey Feedback Meetings: An Evaluation Tool	5. TYPE OF REPORT & PERIOD COVERED: Technical Report
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Patricia A. Pecorella	8. CONTRACT OR GRANT NUMBER(s): N00014-67-A-0181-0013, NR170-719
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Institute for Social Research, University of Michigan	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Organizational Effectiveness Research Programs	12. REPORT DATE: May, 1975 13. NUMBER OF PAGES: ~~82~~ 95
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report) 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Audio Tapes	Data Reduction Techniques	Self-Ratings
Coder Reliability	Evaluation	Supervisory Leadership
Consultant Roles	Problem-Identification	Survey Feedback
Content Analysis	Problem-Solving	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Two coding schemes were developed and subjected to preliminary analyses. The Behavior Classification Scheme (BCS) included eleven behavior and four speaker categories which were used to code each problem-solving input made during a survey feedback meeting. The Summary Rating Scheme (SRS) included several survey questions regarding the feedback meeting which coders answered after listening to an entire meeting. The analyses performed suggest that the two schemes are feasible tools for documenting what transpires during survey feedback sessions.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-014-6601

PRICES SUBJECT TO CHANGE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

TECHNICAL REPORT

May, 1975

CONTENT ANALYSIS OF SURVEY FEEDBACK MEETINGS: AN EVALUATION TOOL

Patricia A. Pecorella

Institute for Social Research
University of Michigan
Ann Arbor, Michigan

This research is funded by the Office of Naval Research, Organizational Effectiveness Research Programs, under Contract No. N00014-67-A-0181-0013, NR170-719.

CONTENT ANALYSIS OF SURVEY FEEDBACK MEETINGS: AN EVALUATION TOOL

During the past several years, one thrust of the Organization Development Research Program* has been to evaluate the effectiveness of organization development (OD) techniques--especially survey feedback. In order to facilitate such evaluations, pre and post measures of organizational functioning and several kinds of criterion data (e.g., productivity, absenteeism) have been collected. Accumulating evidence indicates that survey feedback has been employed successfully in a number of work organizations (Bowers, 1973; Franklin, 1975; Bowers & Hausser, 1975) and suggests that it is a potentially powerful OD technique. This does not imply, however, that survey feedback is unconditionally or uniformly effective. There are also cases where survey feedback efforts have produced, at best, mixed results. Even within one organization, the use of this OD technique may be associated with positive change in some work groups and negligible or negative change in other groups. Results such as these indicate that the effectiveness of an OD technique is not determined solely by elements inherent to the technique itself. Its relative effectiveness is also likely to be influenced by variations in how the basic technique is implemented (Klein, et al., 1971).

---

*The Organization Development Research Program, a part of the Institute for Social Research, is headed by Dr. David G. Bowers. The program previously called the Business, Industry, and Government Program, was formed in 1970.

Two essential elements in survey feedback--problem-identification and problem-solving--are implemented through work group meetings. The success or failure of survey feedback hingers, at least in part, then, on what transpires during these meetings. The fact that the process of group meetings may vary widely--even when general guidelines are provided--points to the desirability of documenting in some detail what occurs during the meetings themselves. Such documentation would make it possible to investigate the effects various styles of implementing survey feedback have on the success of the feedback effort. Based on these findings, specific training objectives and programs could be developed for consultants and managers planning to implement the change technique. In addition, mechanisms for providing intermediate "evaluations" could be built in.

An initial effort was made in the present study to develop and test a documentation method applicable to group feedback meetings. Two coding schemes were developed and applied to a sample of tape-recorded group feedback meetings.* The purpose of this report is to describe the schemes and their reliability and to test, in an exploratory manner, some potential uses of such schemes.

---

*While there were many possible data collection methods (e.g., observation, questionnaires), audio tape recordings of meetings was chosen. This approach offered several advantages: (1) All verbal "raw" data would remain in tact. The data could then be analyzed at a later time in any number of ways. That is, later analysis would not be overly constrained by the initial data collection method. (2) Tape-recording group meetings required little time from the clients for raw data collection. In development efforts time required may be a crucial factor. (3) While group members might be aware of the tape-recorder during meetings, the recorder would probably be less intrusive than a trained observer. (4) Some variation in data collection methods is desirable in itself in the sense that different people's perceptions could be collected in different ways. Audio recordings were a contrast to the survey and interview methods most often used in our development efforts.

Sample and Taping Procedure

The data used in this study were collected during an OD effort which began in December, 1972 in a large business firm. Two departments, employing 324 persons, were the focus of the change effort. In these two departments there were a total of 33 work groups* and each work group had numerous (N > 8) feedback meetings over a period of several months. The sample consists of a subset of these meetings which were tape-recorded.

At the outset of the meetings, the internal consultant requested permission from the group to tape the meetings. Very little resistance to the taping was expressed and no work group refused to be taped. To provide some flexibility concerning what would be taped and under what conditions, work group members were told that they could have the tape recorder turned off at any time. All tapes were regarded as confidential; no one in the company other than the internal consultants had access to them.

For the purpose of testing the coding schemes, only the first two meetings were included in the present sample. Due to some mechanical difficulties in the taping, the tapes from some of these early meetings were not useable. The final sample included at least one meeting for 26 work groups. Nineteen (19) first meetings and eighteen (18) second meetings were successfully taped.

Before turning to the coding schemes and procedures applied to these meetings, a few words about the nature of these meetings seems in order. All of the group feedback meetings had certain characteristics in common. A standardized survey had been administered to all employees in the two

---

*A work group is defined as a supervisor and his or her immediate subordinates.

departments. The summarized survey results fed back to each work group pertained to what the subordinates of that particular group said about their supervisor, their peers, the organizational climate, etc. Each supervisor received the results for his work group and, after reviewing the data with a resource person (i.e., company consultant), shared the results with subordinates in a series of meetings. The purpose of the meetings was to discuss the results and to identify and solve problems indicated by the data. A resource person was "assigned" to each group and attended the meetings to facilitate the understanding and utilization of the survey data. These, then, were the common elements in the survey feedback meetings.

The Coding Schemes

Obviously, tape-recorded discussions provide a massive amount of information which needs to be condensed and summarized in some way if it is to be meaningful. To accomplish this data reduction task, two coding schemes were developed. One--the Behavior Classification Scheme (BCS), defined discrete categories of behaviors. The coders, as they listened to a meeting, recorded in sequence each occurrence of each behavior category. This type of "statement-by statement" coding scheme resembles the classic content analysis procedure employed by Bales since the late 1940's (Bales, 1947; 1950a; 1950b; 1955; 1958; 1970). Bales' interest has been in the study of interpersonal and small group behavior, and thus, his methods and behavior categories were very relevant to the task at hand. Coding schemes which have been applied to classroom interactions (Massialas, et al., 1970; Amidon & Horugh, 1967) and problem-solving

discussions (Mann & Morris, undated) also provided content and structural guidelines for developing the Behavior Classification Scheme.

The strength of a scheme like the BCS lies in the precise, concrete definitions of coding categories that usually accompany them. The importance of providing precise definitions of categories stems from the difficulty of the coding task. Coders must separate an ongoing stream of interaction into discrete "coding units" and simultaneously select the appropriate "label" for each unit. If the categories are not clearly defined, the coders' task quickly deteriorates into a random, hit-and-miss operation. The need for preciseness leads to an emphasis on "behaviors" since it is more difficult to define attitudes, motivations, and feelings in terms of discrete statements. These non-behavioral dimensions of meetings and interactions are often more pervasive and diffuse. Thus, the BCS provides a concise framework for summarizing the behaviors occurring during a feedback meeting meeting.

The second coding scheme--the Summary Rating Scheme (SRS)--included several survey-like questions which the coders answered after listening to an entire meeting. In this case, then, a coding unit consisted of an entire meeting. Rating schemes like these have been employed by on-site observers (Jenkins, et al., 1975) as well as by off-site coders (Alderfer & Lodahl, 1971; Berg, 1972). Schemes like the SRS are quite flexible in terms of the numbers and kinds of dimensions that can be tapped. The coding task is relatively simple, although coder reliability and validity is an issue with this scheme as much as with the BCS.

Many of the questions included in the SRS were drawn from Bowers' conceptualization of consultant roles and consultant debriefing interview schedules employed in an earlier OD effort (Bowers, undated). In addition, some questions were written to assess characteristics comparable to those measured by the BCS so that relationships between the two schemes could be explored.

Because the two schemes differ substantially in their content and format, each will be described separately. The issue of coder reliability is also addressed separately for each scheme since the appropriate techniques for assessing reliability vary according to the format of the coding scheme.

THE SUMMARY RATING SCHEME

The Summary Rating Scheme (SRS) included 54 questions that focused on four aspects of the meetings: the role of the resource person, the role of the supervisor, the role of group members, and the quality and nature of the group discussion as a whole. Most questions were answered on a 5-point extent scale. The entire SRS is included in Appendix A. A sample item and the rating scale are presented below:

Q43. SILENCE: To what extent were there long periods of silence during the meeting?

1. To a very little extent
2. To a little extent
3. To some extent
4. To a great extent
5. To a very great extent

The questions in the SRS were developed in the context of some a priori assumptions regarding the appropriate roles for the various people. For example, the role of the resource person (RP) was seen as one of facilitating the discussion without leading the meeting. Thus, questions about the resource person focused on the ways in which this facilitating role was accomplished, the resource person's activity level, the clarity of the resource person's statements, and the response of the group to the resource person's inputs. Three alternative facilitator styles were measured. These styles are referred to as information resource, catalyst, and interpersonal confronter.

The extent to which the resource person acted as an information resource was measured by a single, straightforward item. This style was defined for the coders as one in which the resource person provides information regarding the meaning of the group's survey results, the goals of the Survey-Guided Development effort, and activities that might take place as a part of the effort. The extent to which the RP acted as a catalyst was defined as the extent to which the RP made inputs designed to promote indirectly the problem-solving process. This style was measured by two items--one assessing whether the RP affected the process of the meeting, and one regarding the extent to which the RP stimulated action steps. The extent to which the RP acted as an interpersonal confronter was assessed by three questions. Two of these questions measured the extent to which the RP made statements about how group members related to each other and actively confronted them. One additional item assessed the extent to which the RP intervened in emotional encounters between group members.

The RP's overall activity level was measured by items referring to dominating and directive behavior, the extent to which the RP made more inputs than the group desired, and a reversed item that measured the extent to which the RP remained silent except to answer questions. The degree of opposition of group members to the RP's interventions was also assessed.

Two aspects of the supervisor's role as group leader were of interest: (1) the extent to which the supervisor exhibited a variety of leadership behaviors and was competent in leading the discussion, and (2) the extent to which the supervisor prevented or suppressed participation by group members.

The role played by group members was assessed for all subordinates as a group rather than for individual people. The variables of interest were: (1) the extent to which several group members actively participated in the discussion, and (2) the extent to which a few dominated the discussion.

Finally, there were several questions that focused on how the discussion, as a whole, progressed. The areas tapped were varied and include the general types of problems dealt with, how productive and systematic the meeting was, the degree of frustration, defensiveness, resistance, and commitment exhibited, the extent to which people were listening to each other, and the extent to which conflicting messages were being communicated.

The coding procedure involved listening to an entire tape and then answering the questions with reference to the meeting as a whole. The author held one session with the coders to train them in using the SRS. During the session, each question was read aloud, paraphrased by the coders, and explained in more detail when there were ambiguities.

Sample portions of tapes were played after which the coders and the author answered independently all 54 questions. The ratings were compared, with the investigator's ratings serving as the standard. There were only minor discrepancies in ratings and these were discussed and resolved. This training session lasted about three hours. It was decided that no further formal training was needed, although if questions regarding the scheme arose, the coders were asked to consult with the investigator.

Summary Rating Scheme: Coder Reliability and Validity

A major purpose of this investigation was to develop and test methods of assessing what transpires during survey feedback meetings. The methodology chosen in this case was behavior coding of audio tapes. Central issues affecting the utility of this method are (1) the extent to which coders rate the same behaviors in the same way, and (2) the extent to which coders rate behaviors correctly (i.e., the way an expert judge would rate them). These issues are essentially issues of reliability and validity: How confident may one be that the ratings reflect what actually occurred during the meeting as opposed to reflecting the individual orientations of the coders?

The first step in answering this question usually involves verifying the units that were coded. Unit reliability is not an issue in this case, however, because the Summary Rating Scheme defined a unit as the audio tape of a meeting. Thus there was, for all practical purposes, zero probability of error in unit definition.

The second step is to assess the extent to which coders' ratings are congruent with each other (reliability) and with a "standard" (validity). In this case, the investigator served as the reliability check and the standard simultaneously. A subset of tapes were coded independently by the investigator as well as being rated by one other coder. The task then was to assess the extent to which the ratings of the coder and the investigator were congruent.

The two most common tests of inter-coder agreement are the correlation coefficient and a simple percentage of agreement coefficient. Each of these methods, however, has weaknesses. A correlation coefficient is a measure of association rather than agreement. Thus, one coder might give a series of dimensions ratings of 1, 2, 2, 3, 3, while the check coder gives ratings of 3, 4, 4, 5, 5. In this example dimensions are rated relatively higher or lower to the same extent by both coders. Thus, the correlation coefficient would be +1.0, which represents perfect association, even though the two coders are not giving units the same ratings. The problem with percentage agreement statistic is that it assesses coder agreement without taking account of the fact that some degree of congruence is expected by mere chance alone. Thus, what is needed is a test of agreement which assesses the degree of congruence greater than that expected by chance. Cohen's weighted kappa (k_w) (Cohen, 1968) is one such test and was used to assess coder agreement with the standard for the SRS. An additional feature of Cohen's test is that partial credit may be

given for two close but non-identical ratings through assigning weights to the off-diagonal (diagonal = perfect agreement) cells. The formula for weighted kappa is:

$$k_w = \frac{\Sigma^w_{ij} f_{oij} - \Sigma^w_{ij} f_{cij}}{w_{max} N - \Sigma^w_{ij} f_{cij}}$$

where:

Σ = sum
w_{max} = the maximum weight
w_{ij} = the weight assigned to a given cell
f_{oij} = the observed frequency in a given cell
f_{cij} = the frequency expected by chance in a given cell = row total/N X column total
N = total frequency

The value of k_w may range from -1 to +1; however, maximum values require having identical row and column totals. In practice, then, the maximum value of k_w is somewhat less than 1.

In the present study, partial credit (2/3 credit) was given for ratings that were discrepant by 1 scale point. No credit was given when discrepancies between the coder's and the investigator's ratings were greater than 1 scale point. Thus, the following weighting scheme was applied:

Investigator Ratings (5-point scale)

Coder Ratings (5-point scale)

	1	2	3	4	5
1	3	2	0	0	0
2	2	3	2	0	0
3	0	2	3	2	0
4	0	0	2	3	2
5	0	0	0	2	3

= perfect agreement

Seventeen tapes were coded by both the investigator and a coder. It was necessary to compute global reliabilities rather than computing a k_w for each question since the number of tapes which were coded by both was small (17). Thus, the reliabilities of sets of questions were assessed without regard to the content of the specific items. The k_w coefficients reflect, therefore, inter-rater agreement in general rather than coder reliability on particular questions.

For the purpose of computing k_w, the 54 questions were put into three a priori categories:

Category I: Questions that required mental summing and averaging of observable behaviors. (N = 27 questions)

Category II: Questions that required the coders to make judgements about the quality of a behavior or contribution. (N = 21 questions)

Category III: Questions that required the coder to make inferences about the attitudes or desires of group participants. (N = 6 questions)

Questions in Category I were the most straightforward and objective questions, while questions in Category III were more difficult and subjective in nature. Based on the assumption that coders are likely to view observable behaviors more similarly than they interpret attitudes and desires, it was predicted that Category I would have the highest k_w value and Category III would have the lowest. The category into which each question was placed is indicated in Appendix B. In order to give the reader a feel for the three categories, however, sample items from each category are listed below:

Sample Items from each A Priori Category

Category I:	Q2.	To what extent did the resource person make statements about how group members were relating to each other?
	Q9.	To what extent did the resource person remain silent except when responding to questions directed toward her/him?
Category II:	Q1.	To what extent did the resource person make comments which affected the process of the discussion?
	Q8.	To what extent were the resource person's statements easy to understand? (i.e., clarity in meaning)
Category III:	Q4.	To what extent did the resource person make more inputs than the group seemed to want?
	Q16.	To what extent were double messages being sent? That is, to what extent were there discrepancies between a participants verbal message and his emotional message?

A weighted kappa coefficient was calculated separately for each category.

Table 1 presents the results. The k_w coefficients ranged from .54 to .43. As was predicted, Category I had the highest reliability and Category III the lowest although the variation was not substantial. These reliability coefficients are in line with those obtained by other researchers using k_w. Jenkins, et al., (1975) when computing the inter-rater reliability on several job and personality characteristics, got k_w coefficients ranging from -.67 to -.26, with 27 of 59 coefficients being smaller than $|.33|$. Thus, the SRS is judged to be acceptable, reliable, and valid.

Table 1

WEIGHTED KAPPA (k_w) COEFFICIENTS FOR THREE A PRIORI CATEGORIES OF THE SUBJECTIVE CODER ANALYSIS: RELIABILITY AND VALIDITY ESTIMATES

Category	K_w
I	.54
II	.49
III	.43

Summary Rating Scheme: Index Creation

As stated previously, SRS items focused on selected aspects of the roles played by the resource person, supervisor, and group members, and on the nature of the group discussion as a whole. This framework influenced the index creation process. In addition, however, a hierarchical cluster analysis was performed using the data from one meeting per group (N=26).*

Thirteen (13) indices, including between 2 and 6 items each, were created. These indices and their component items are listed in Table 2. Twelve (out of the 54) SRS items were not included in any index but were retained as single items. These items are also listed in Table 2.

Alpha coefficients, a measure of internal consistency or scale reliability, were computed for each of these indices. The results are shown in Table 3. The alpha's ranged from .63 to .94. Thus, the internal consistency of the indices is quite acceptable.

Table 4 shows the intercorrelations among the indices. Some moderate correlations were expected because all the SRS items shared methods variance and because the sets of behaviors reflected by the indices are probably related to some degree "in reality." At the same time, however, very high intercorrelations would indicate that the indices measured the same or very similar concepts. As Table 4 shows, the correlations ranged from .00 to .77. The median intercorrelation was .20. Thus, the correlations varied substantially in size. No inter-scale correlation was larger--and

*If a group had data for two meetings, data from the first meeting were used.

Table 2

SRS INDICES AND COMPONENT ITEMS

Index	Component Items - "To what extent...
SRS1: RP Makes Impact	Q1...did the RP make comments which affected the process of the discussion?"
	Q6...was the RP active in making diagnostic comments and/or stimulating action steps
SRS2: RP Overly Active/Directive	Q4...did the RP make more inputs than the group seemed to want?"
	Q9...did the RP remain silent except when responding to questions directed toward him/her?" (Reversed item)
	Q50...did the RP dominate the discussion?"
	Q53...did the RP actively direct the discussion?"
SRS3: RP Confronts Group Members	Q2...did the RP make statements about how group members were relating to each other?"
	Q7...did the RP actively confront group members?"
SRS4: RP Opposed by Group	Q10...did conflict exist between the RP?"
	Q11...When the RP intervened, to what extent was encountered in response to the intervention?"
	Q13...was the RP attacked by the group for suggesting changes?"
SRS5: Supervisory Leadership	Q15...was the supervisor competent in leading the discussion?"
	Q43...did the supervisor seem open to the opinions and ideas of his/her subordinates?"
	Q44...did the supervisor encourage group members to work together as a team?"
	Q45...did the supervisor help to remove roadblocks to solving problems?"
	Q46...did the supervisor emphasize goals (work, change objectives, etc.)?"
	Q51...did the supervisor actively direct the discussion?"

Index	Component Items - "To what extent...
SRS6: Sup. Prevents Group Discussion	Q47...did the supervisor suppress group discussion through his/her attitude or actions?"
	Q48...did the supervisor dominate the discussion?"
SRS7: Group Participation	Q38...did group members participate in the discussion?"
	Q39...did group members interact with each other (rather than merely responding to the supervisor)?"
	Q40...were group members willing to talk about problems?"
	Q52...did group members actively direct the discussion?"
SRS8: Group Not Defensive	Q17...were members (including the supervisor) honest and candid about themselves?"
	Q18...did individuals take an "I'm the-one-who-has-to-change" attitude?"
SRS9: Productive Meeting	Q30...were maintenance inputs provided by the RP, supervisor, and group members?"
	Q31...were content inputs provided?"
	Q32...were real problems identified?"
	Q33...were there attempts to solve problems which were identified?"
	Q34...were problems actually solved?"
	Q37...was the group interaction of "high quality?"
SRS10: Discussion Unsystematic	Q20...was the problem-solving sequence (the content inputs) followed followed systematically?" (Reversed item)
	Q21...did the discussion get side-tracked on to inconsequential topics?"

Index	Component Items - "To what extent...
SRS11: Change Objectives Opposed	Q26...were change objectives of the program clearly explained to group members?"
	Q27...was there disagreement on change objectives?"
	Q28...was there resistance to the change objectives?"
SRS12: Problems Outside Control	Q19...was a "change-the-other-guy" attitude in evidence?"
	Q22d...(were) problems related to constraints created by the situation?"
	Q35...were the problems discussed within the control of the work group and supervisor?" (Reversed item)
	Q36...were the problems discussed outside the control of the work group and supervisor?"
SRS13: Group Not Listening	Q14...did group members (including the supervisor) misunderstand each other?"
	Q42...were people "battling" for air-time?" i.e., to what extent were people interrupting each other and talking all at once in order to get their opinions voiced?"
	Single Items - "To what extent...
	Q3...did the group request inputs from the RP?"
	Q5...did the resource person act as an information resource?"
	Q8...were the RP's statements easy to understand?"
	Q12...did the RP intervene in emotional encounters between others at the meeting?"
	Q16...were double messages being sent? That is, to what extent were there discrepancies between a participant's verbal message and his emotional message?"

	Single Items - "To what extent...
	Q50...did a few group members dominate the discussion?"
	Q22a...(were) problems related to feelings or affect or values?"
	Q22b...(were problems related to the lack, or inaccuracy, of information?"
	Q22c...(were) problems related to the lack of needed skills?"
	Q25...was there consensus on and commitment to the solutions advanced?"
	Q29...how much frustration was in evidence?"
	Q41...were there long periods of silence during the meeting?"

Table 3

ALPHA COEFFICIENTS FOR SRS INDICES

		Items	Alpha
#1:	RP/Makes impact	2	.88
#2:	RP/overly active/directive	4	.94
#3:	RP/confronts group members	2	.86
#4:	RP/opposed by group	3	.63
#5:	Sup leadership	6	.80
#6:	Sup prevents group discussion	2	.70
#7:	Group participates	4	.89
#8:	Gropp not self-protective	2	.69
#9:	Productive meeting	6	.89
#10:	Discussion unsystematic	2	.73
#11:	Change objectives opposed	3	.83
#12:	Problems outside control	4	.83
#13:	Group not listening	2	.68

Table 4

INTERCORRELATIONS OF SRS INDICES

SRS												
2	.70											
3	.33	.16										
4	.42	.38	.47									
5	-.05	-.30	-.13	-.13								
6	.19	-.04	.06	.31	.10							
7	-.10	-.31	.00	.03	.31	-.06						
8	.02	-.23	.05	-.08	.77	.20	.39					
9	.07	-.28	-.09	-.13	.62	.30	.58	.57				
10	.16	.31	-.13	.02	.10	.09	.38	.35	.26			
11	-.01	-.25	.41	.01	.33	-.14	.29	.43	.18	-.15		
12	-.01	.12	-.24	-.23	-.05	-.05	.22	.02	.33	.48	-.13	
13	.12	-.06	-.10	.02	.38	.20	.67	.49	.62	.55	.35	.54
	1	2	3	4	5	6	7	8	9	10	11	12

SRS1 = RP Makes Impact
2 = RP Overly Active/Directive
3 = RP Confronts Group Members
4 = RP Opposed by Groups
5 = Supervisory Leadership
6 = Sup. Prevents Group Discussion
7 = Group Participation
8 = Group Not Defensive
9 = Productive Meeting
10 = Discussion Unsystematic
11 = Change Objectives Opposed
12 = Problems Outside Control
13 = Group Not Listening

most were substantially smaller--than the alphas for the respective scales, however. This is a sign of good discriminant validity--that is, that the SRS indices measured several different dimensions of the meeting.

THE BEHAVIOR CLASSIFICATION SCHEME

The Behavior Classification Scheme (BCS), the second coding scheme developed in this study, was used to classify the verbal interactions during survey feedback meetings. The interactions were coded directly from the audio tapes. Coders were instructed to listen to the entire tape, apply the Summary Rating Scheme, and then go through the tape again using the BCS. This sequence was adopted in order to give the coders some familiarity with the general flow of the meeting before they began their behavior-by-behavior coding of the often complex, fast-moving verbal interactions.

The entire coding scheme in its original form is included in Appendix A. A summary of the original categories and sub-categories in the scheme is presented in Table 5. The categories bear a marked resemblance to the functions or roles that have been described as important aspects of group process by many researchers and practitioners. There were 37 basic categories: 16 were task-related, 16 were maintenance related, and 5 were referred to as miscellaneous inputs. Two basic categories were sub-divided into more specific categories, making a total of 42. All 42 categories were defined in terms of the problem-solving behavior of any speaker; no single category was restricted to, for example, resource person statements or group member statements.

A coding unit in the BCS was defined as a problem-solving input provided by a given speaker during a survey feedback meeting. A unit consisted of a single and complete task, maintenance (i.e., process), or miscellaneous input as defined by the behavior categories, regardless of the time required to make the input. In addition, every time the "speaker" (resource person, supervisor, group member) changed, a new unit was coded.

Categories 10, 15, 30, 31, 35, and 50 to 53 were "request" categories; that is, they defined information-seeking, question-asking behaviors. Categories 20, 25, 40 to 44, and 60 to 63 were "provision" categories paralleling the request categories in content. The provision categories, however, describe information-giving, question-answering behaviors. Categories 70 to 79 were considered functional inputs that would help to maintain or promote constructive discussion. Categories 82 to 87 were considered dysfunctional inputs; that is, they were described as inputs that might hinder group discussion. The behaviors described by categories 95 to 98 were also related to maintaining group discussion, but were not classified as functional or dysfunctional inputs. Finally, category 99 was used to indicate that the coders could not hear what was being said and was considered the equivalent of a missing data code in, for example, survey data.

There were, in addition, 5 notations for speakers:

1 = resource person
2 = supervisor
3 = work group member
4 = general response
9 = unidentifiable speaker (missing data equivalent)

Table 5

SUMMARY OF CATEGORIES IN THE BEHAVIOR CLASSIFICATION SYSTEM

TYPE OF INPUT

	REQUEST	PROVISION
CONTENT INPUTS	10. Exposition 15. Progress reporting 30. Problem definition and clarification 31. Problem in whose control 35. Setting priorities 50. Identifying solutions 51. Evaluating solutions 52. Active choice of solutions 53. Implementing solutions	20. Exposition 25. Progress reporting 40. Problem definition and clarification 41. Problem within group's control 42. Problem outside group's control 43. Task-related problems 44. Interpersonal conflict problem 45. Setting priorities 60. Identifying solutions 61. Evaluating solutions 62. Active choice of solutions 63. Implementing solutions
	FUNCTIONAL	DYSFUNCTIONAL
MAINTENANCE INPUTS	70. Encouragement of participation 71. Reinforcement/acceptance 72. Positive perception 73. Agreement 74. Disagreement 75. Separation 76. Consensus testing 77. Conflict resolution 78. Summarizing 79. Process comments/checking accuracy	82. Negative perception 83. Negative response 84. Non-productive response 85. Non-response to leads 86. Fragmented discussion 87. Suppression
MISC. INPUTS	95. Seeking approval 96. Evidence of misunderstanding 97. Laughter 98. Miscellaneous 99. Unintelligible/NA 00.	

A. Request of Task Inputs

10. EXPOSITION: The speaker requests statements which provide general information about or reactions to (a) the development effort, survey methodology, or meaning of the indices, (b) the purpose of the data feedback session(s), or (c) the score on an index or item.

15. PROGRESS REPORTING: The speaker requests statements which provide information about (a) which action steps have been taken to solve a problem or (b) whether the action steps taken have solved the problem they were intended to solve.

30. PROBLEM DEFINITION AND CLARIFICATION: The speaker requests statements which (a) identify specific, concrete problems indicated by a particular index or item score, (b) clarify a previous statement, or (c) provide a concrete example of a general problem.

31. Problem Within Whose Control: The speaker requests statements which identify the problem as one which that particular work group can solve or as one which another group or department must solve.

35. SETTING PRIORITIES: The speaker requests statements which indicate the order in which the problems defined should be solved.

50. IDENTIFYING POSSIBLE SOLUTIONS: The speaker requests statements which identify possible solutions to a problem in terms of the concrete steps which must be taken.

51. EVALUATING PROPOSED SOLUTIONS: The speaker requests statements which evaluate the feasibility, attractiveness, or utility of the suggestions for solving a problem.

52. ACTIVE CHOICE OF SOLUTION: The speaker requests statements which indicate which of the proposed solutions will be implemented.

5[illegible]. IMPLEMENTING THE SOLUTION: The speaker requests statements which identify a particular person as responsible for taking specific action steps.

B. PROVISION OF CONTENT INPUTS

20. EXPOSITION: The speaker makes statements which provide general information about or reactions to (a) the development effort, survey methodology, or the meaning of indices, (b) the purpose of the data feedback session(s), or (c) the score on an index or item.

25. PROGRESS REPORTING: The speaker makes statements which provide information about (a) which action steps have been taken to solve a problem or (b) whether the action steps taken have solved the problem they were intended to solve.

40. PROBLEM DEFINITION AND CLARIFICATION: The speaker makes statements which (a) identify specific, concrete problems indicated by a particular item or index score, (b) clarify a previous statement, or (c) provide a concrete example of a general problem.

41. Problem Within Group's Control: The speaker makes statements which identify the problem as one which that particular work group can solve.

42. Problem Outside the Group's Control: Speaker makes statements which identify the problem as one which someone outside that particular work group must solve.

43. Task-Related Problem: The speaker makes statements indicating that the problem is a task-related one as opposed to one centering around interpersonal conflicts.

44. Problem Centers Around Interpersonal Conflict: The speaker makes statements indicating that the problem centers around interpersonal conflict as opposed to being a very task-related problem.

45. SETTING PRIORITIES: The speaker makes statements which indicate the order in which the problems defined should be solved.

60. IDENTIFYING POSSIBLE SOLUTIONS: The speaker makes statements which identify possible solutions to a problem in terms of the concrete steps which must be taken

61. EVALUATING PROPOSED SOLUTIONS: The speaker makes statements which evaluate the feasibility, attractiveness, or utility of the suggestions for solving a problem.

62. ACTIVE CHOICE OF SOLUTION: The speaker makes statements which indicate which of the proposed solutions will be implemented.

63. IMPLEMENTING THE SOLUTION: The speaker makes statements which identify a particular person as responsible for taking specific action steps.

C. PROVISION OF MAINTENANCE INPUTS

Functional Inputs:

70. ACTIVE ENCOURAGEMENT OF PARTICIPATION: The speaker encourages others to ask questions, offer opinions, and discuss issues by explicitly inviting such participation.

71. REINFORCEMENT/ACCEPTANCE: The speaker makes statements which indicate that the individual should continue his behavior though not necessarily indicating agreement with content.

72. POSITIVE PERCEPTION OF OTHERS: Speaker makes statements which indicate a positive view of some other person(s). Use this category only when the unit is not coded as a content input.

73. EXPRESSING AGREEMENT: The speaker makes statements which clearly indicate agreement with a statement or idea.

74. EXPRESSING DISAGREEMENT: The speaker makes statements which indicate disagreement with an idea or statement.

75. SEPARATING IDEA-SEEKING FROM IDEA-EVALUATION: The speaker makes statements requesting that a problem be defined or suggestions for solving problems be presented without anyone evaluating their feasibility, utility, or attractiveness.

76. CONSENSUS-TESTING: The speaker requests statements indicating the degree of agreement with an issue or decision.

77. CONFLICT RESOLUTION: The speaker makes statements which encourage two or more people to stop disagreeing with each other so vehemently. The speaker's statements indicate that he/she is seeking the source of the conflict or trying to resolve it, not just trying to cool it out.

78. SUMMARIZING: The speaker makes statements reviewing the discussion. The speaker is doing more than paraphrasing another speaker; he/she is also integrating previous discussion.

79. PROCESS COMMENTS/CHECKING ACCURACY: The speaker makes statements which help to guide the discussion, get the group back on the subject, keep the group focused on the data, or check the accuracy or clarity of some statement (speaker's own or someone else's).

Dysfunctional Inputs:

82. NEGATIVE PERCEPTION OF OTHERS: Speaker makes statements which indicate a negative view of some other person(s). It may take the form of an accusation. Use this category only when the unit is not coded as a content input.

83. NEGATIVE RESPONSE: The speaker indicates an inability or unwillingness to respond to a request or perform a task.

84. NON-PRODUCTIVE RESPONSE: The speaker makes irrelevant or disruptive statements.

85. NON-RESPONSE TO LEADS: Period of silence following a question or request, ended by (1) speaker clarifying lead, or (2) same or different speaker going off in a different direction. "Thinking time" of short duration (5 seconds or so) should not be coded as non-response. This category is trying to pick up unwillingness to respond, or inability to respond because of the nature of the question or request. Code speaker as "general response", unless a specific speaker was addressed. Then code the speaker's number who did not respond

86. FRAGMENTED DISCUSSION: A period which cannot be categorized because the statement(s) cannot be understood or cannot be separated.

87. SUPPRESSION: The speaker makes statements which inhibit or squash discussion by indicating that certain topics are not to be discussed because of his/her personal views.

D. MISCELLANEOUS INPUTS

95. SEEKING APPROVAL: The speaker requests statements which indicate that his behavior is acceptable to others.

96. EVIDENCE OF MISUNDERSTANDING: The speaker calls attention to a misunderstanding. (Checking understanding should be coded as a process comment).

97. LAUGHTER: Laughter which interrupts the flow of conversation.

98. MISCELLANEOUS: The speaker makes statements which cannot be classified under any of the other categories.

99. UNINTELLIGIBLE: A statement cannot be understood because it is inaudible or drowned out by background noise.

For each unit identified by the coders, both the speaker and the behavior code notations were recorded.* The voices of the resource person and supervisor were identified for the coders at the start of each tape by the present investigator.

Training sessions of various kinds were held with professional coders at the University of Michigan's Survey Research Center over a time span of four weeks. Four coders and one check coder/coding coordinator were involved. The present investigator also served at times as a trainer and check-coder. The training itself was a two-step process. First the investigator and the project coding coordinator worked together to refine the scheme. During this time, the coordinator became very familiar both with the BCS and with this investigator's definitions and interpretations of the categories. One sample tape was coded jointly by the coordinator and investigator, and portions of a second tape were coded independently. The two sets of codes were then compared and disagreements were resolved. Second, the coders assigned to the project, the coordinator, and the coders. Once again, the scheme was discussed in detail, a tape was coded jointly, and then portions of other tapes were coded independently and everyone's ratings compared. When coding discrepancies began to occur infrequently, the formal training was terminated. Coders were instructed to go to either the coding coordinator or the research investigator when questions regarding the scheme arose. Close, frequent contact (usually daily) was maintained with the coders throughout the time tapes were being coded.

---

*In addition to a more detailed explanation of the categories, Appendix A includes examples and guidelines for coding the verbal interactions.

Behavior Classification Scheme: Coder Reliability

The Behavior Classification Scheme with its 42 discrete categories and sub-categories, was a complex scheme to apply to the audio tapes. First, coders had to make decisions concerning where one coding unit ended and another began. This decision was simple when speakers changed since every time a new speaker began, a new unit was coded. There were also times, however, when the same person spoke for a fair length of time (e.g., 15 minutes). In such cases, coders were instructed to use the behavior category definitions in deciding when a problem-solving input was completed. The inputs made during each meeting were recorded in sequence rather than merely keeping frequency counts of the behaviors occurring in order to preserve the flow of the discussion.

In order to check the raters' consistency in defining and labelling units, a fairly rigorous "check-coding" procedure was established. Coders were instructed to write down the key parts of statements in addition to the unit labels for two units out of every sequence of 21 units coded. These written keys made it possible to match coded units with the original tape-recorded interactions.

The coding coordinator served as the standard of comparison for the BCS. The coders applied the BCS to 20 tapes in this sample--16 first meetings and 4 second meetings. Portions of 17 tapes which the coders worked through were also coded independently by the coding coordinator. For each check-coded tape, the coordinator's ratings were compared to the coder's ratings of the same portion of the meeting. The coder and check-coder met to discuss their ratings and reached consensus where

discrepancies were present. This discussion served as additional training for the coders. The consensus codes were recorded but were not used in testing reliability since the goal in checking the reliability was to see how congruent the two initial, independent ratings were.

The number of units coded for the meetings ranged from 189 to 882 units. The number of units per meeting check-coded (i.e., the number of units per tape checked by the coding coordinator) ranged from 49 to 63 units. Thus, portions of 85% of the tapes in the sample were check-coded. In terms of the percentage of units which were check-coded, one may assume that a minimum of 7% (49/882) and a maximum of 33% (63/189) were coded independently by two people.

When the two sets of ratings for the 17 tapes were initially displayed, it was apparent that some refinement and condensing of the coding scheme was needed. Many of the categories were used very infrequently (namely, categories 15, 25, 31, 41 to 44, 35, 45, 61 to 63, 51 to 53, 74 to 78, 82 to 85, 87, and 96). In addition, some confusion appeared to exist about the distinctions among some categories. For example, categories 71 (reinforcement) and 73 (agreement) were often used interchangeably--one rater would label a unit 71, while the other rater would use category 73. Thus, the coding scheme was revised by combining each of the infrequently used categories with other related, similar categories and by collapsing categories that were used interchangeably.

The Behavior Classification System, after this revision, included twelve (instead of 42) categories. The categories were as follows:

BEHAVIOR CLASSIFICATION SCHEME: REVISED

TYPE OF INPUT

	REQUEST	PROVISION
CONTENT INPUTS	10. Problem-identification: Includes categories 10, 15, 30, 31, from original scheme 50. Problem solution: Includes categories 35, 50 to 53, from original scheme	20. Problem-identification: Includes categories 20, 25, 40 to 44, from original scheme 60. Problem solution: Includes categories 45, 60 to 63, from original scheme
	FUNCTIONAL	DYSFUNCTIONAL
MAINTENANCE	71. Reinforcement: Includes categories 70, 71, 73, from original scheme 72. Positive Process: Includes categories 72, 74 to 78, from original scheme 79. Process comments: Same as original	82. Negative Process: Includes categories 82 to 85, 87, 95, 96, from original scheme 86. Fragmented discussion: Same as original
MISC	97. Laughter: Same as original 98. Miscellaneous comments: Same as original 99. Unintelligible comments (Missing Data): Same as original	

A Chi Square (χ^2) test was performed using the revised BCS to check whether the coders and the check coder used the categories with about the same frequency. Thus, the basic question was whether the frequency distributions for the two sets of ratings were quite similar. The χ^2 coefficient was 6.07 ($p<.90$), which indicated that the distributions were very similar. In order to express this in terms of "coder-agreement" or reliability, a modified Scott coefficient, π_m, developed by Flanders (1960) was computed.* Scott's formula is:

$$\pi = \frac{P_o - P_e}{100 - P_e}$$

where: P_o = Percent actual agreement

P_e = Percent agreement expected by chance (which is found by squaring the percentage frequency in each category and summing these over all categories.)

The method developed by Flanders involves the following calculations:

1. Tally the frequency for each category for each coder.
2. Compute the percentage of tallies in each category for each coder.
3. Compute the percentage difference between coders for each category. The sum of these differences is the percent disagreement and 100 - % disagreement = P_o.
4. Compute the average percent falling in each category and square it. The sum of these average percent figures is the estimate of P_e.

Thus, the modified formula becomes:

$$\pi_m = \frac{(100 - \%\ \text{disagreement}) - (\text{average}\ \%)^2}{100 - (\text{average}\ \%)^2}$$

*Scott's formula (Scott, 1955) is usually used to assess unit-by-unit agreement. The modified version developed by Flanders may be used to assess agreement in terms of frequency of category use.

Flanders (1960) stated that a π_m of 0.85 or higher demonstrated a reasonable level of reliability, based on his investigation of the confidence limits.

In the present case, the value of π_m was .92. Thus, the coding scheme with the second set of refinements discussed above passed the first test of reliability. Since category 99 (unintelligible comments, i.e., missing data) contributed no content information to the scheme, and since the X^2 and π_m values indicated that category 99 was used with approximately equal frequency by any two raters, this category was eliminated from further analyses.

Three questions regarding the reliability of the coding scheme remained to be answered:

1. How much agreement was there between the check-coder and the coders in how units were defined?
2. How much coder agreement was there in speaker identification?
3. To what extent did the two raters agree in assigning behavior categories to units?

The two sets of ratings (i.e., check-coder vs. coders) were compared in various ways in order to address each of these issues. The revised BCS was used in each case.

Unit agreement. A total of 1,053 units were identified by either the check-coder or coders on the check-coded portions of the 17 tapes. Of these, 722 units were identified by both raters, constituting 68.6% agreement on unit definition. Among the 331 units over which there was disagreement, 132 were identified by the coders and not by the check-coder, and 199 units were identified by the check-coder and not by the coders. Thus, the check-coder tended to identify more units than did the coders, although each rater identified some units not identified by

the other. Overall, it appears that the two sets of raters saw the various kinds of behaviors defined by the final form of the scheme with the same frequency (as demonstrated by the χ^2 and π_m coefficients) but broke statements down into units somewhat differently. A unit agreement of 68.6% is not overwhelmingly high; but neither is it disappointing, or unacceptably low given the complexity of the rating task, and the similarity of the frequency distributions of the two raters.

Speaker Identification. The next step was to assess how reliable the two sets of raters were in identifying the voices of the resource persons, supervisors, and group members. Cohen's unweighted kappa (K) (1960) was used to assess the reliability of speaker identification.* The formula is $K = \frac{f_o - f_o}{N - f_c}$ and has the same rationale and chance-agreement correction as K_w. The difference between K and K_w is that K is used for assessing the reliability of nominally-scales categories and thus, involves no weighting scheme.**

As stated above, there were five speaker notations, one for the resource person (1), the supervisor (2), any group member (3), a general response (4), and a notation for an unidentifiable speaker (9). There was only one instance in which any rater could not identify the speaker's voice when the statement itself was audible.

---

*Scott's formula (Scott, 1955) was not used because it assumes that the average frequency across judges for all categories will be equal, an assumption the present investigator was unwilling to make.

**See pp. 10-11 for a more complete discussion of k_w as an estimate of coder reliability.

Taking the 722 units on which there was rater agreement on the existence of a unit per se, the Kappa's value was .95. This coefficient demonstrates very high reliability and indicates that the speaker was correctly identified almost without exception.

Unit Labelling. Finally, the reliability of the revised Behavior Classification Scheme was checked in terms of the extent to which the two independent raters assigned the same behavior code to the same unit--i.e., the degree of unit-by-unit agreement in labelling. Coder agreement in unit labelling was assessed separate from coder agreement in speaker identification because identifying the speaker and the behavior were, in fact, two independent operations even though both were recorded for each unit. Once again, Cohen's K (1960) was used to estimate reliability. The following guidelines were established regarding category 99 which, as stated above, was eliminated from the final form of the coding scheme:

For a given unit:

> If the check-coder had a 99, the coder labelled the unit with something other than 99, and the consensus code was the same as the coder's, this was tallied as an instance of agreement. If the coder had a 99, the check-coder used another category, and if the consensus code was the same as the check-coder's, the unit was not counted. If both the check-coder and the coder had a 99, the unit was not counted.

The total number of units used in computing Kappa was 722. The value of Kappa was .726, which represents a very acceptable level of chance-corrected coder agreement.

Summary: Overall Reliability of BCS

The original Behavior Classification Scheme with its 42 categories, proved to be too complex for use in coding audio tapes. Thus, on the basis of early analyses regarding the similarity of the frequency distributions of the two sets of raters, the scheme was condensed by combining sets of related behavior categories. The refined scheme included 11 behavior categories and five speaker notations. Although the refined version would yield less information than its original counterpart, it still provided a means for extracting valuable information from the survey-feedback meetings.

The refined Behavior Classification System demonstrated encouraging reliability in terms of the frequency distributions of the behavior categories, unit definition, unit labelling, and speaker identification. The level of reliability is comparable to other schemes for coding verbal interactions where transcriptions of tapes were used rather than the audio tapes themselves (e.g., Massialas, 1970). This is promising since transcribing tapes is a very time-consuming and expensive venture.

Behavior Classification Scheme: Behavior Indices

Initially the BCS data for each meeting were streams of three-digit codes with the first digit representing the speaker and the last two digits representing the verbal behavior performed. In order to organize and summarize these data, two data management operations were performed. First, the frequency of each input for each meeting was calculated. Three "classes" of variables were relevant: (1) the frequency of inputs by each speaker independent of the behavior performed, (2) the frequency of each behavior independent of the speaker performing it, and (3) the frequency of each behavior performed by each speaker. Calculating all three sets of frequencies maintained considerable flexibility in the types of behavior indices that could be created.

Raw frequencies could not be used to construct indices, however. The meetings varied substantially in length and the frequencies varied accordingly. Thus, raw frequencies could not be compared in a meaningful way across meetings. The second data management step, therefore, was to convert the raw frequencies to proportions.

As stated previously, the basic behavioral dimensions reflected by the BCS were the content, process, and mode of discussion.* Each dimension included two components: content inputs included problem-identification and problem-solving behaviors; process inputs included functional and dysfunctional behaviors; the mode of discussion included

---

*The miscellaneous category might be considered a fourth "dimension".

information-seeking and information-providing inputs. The "behavior indices" created quantify these dimensions and their components by converting the frequencies of these behaviors to ratios. In all, 27 behavior indices were created. They are operationally defined below.

CONTENT INDICES:

Problem-Identification Index (PI) -- the proportion of problem-identification inputs made. The PI index was computed for each speaker (i.e., resource person, supervisor, group members) by summing the frequencies for categories 10 + 20 and dividing by the total number of inputs. For example, the formula for the resource person was: $\frac{\Sigma 110 + \Sigma 120}{\Sigma 1}$

Problem-Solving Index (PS) -- the proportion of problem-solving inputs made. The PS index was computed for each speaker by summing the frequencies for categories 50 + 60 and dividing by the total number of inputs. For example, the formula for the resource person was: $\frac{\Sigma 150 + \Sigma 160}{\Sigma 1}$

Problem-Solving/Content Index (PS/C) -- the proportion of all content inputs which were problem-solving inputs. This index was calculated for the meeting as a whole. The formula was: $\frac{\Sigma 50 + \Sigma 60}{\Sigma 10 + \Sigma 20 + \Sigma 50 + \Sigma 60}$.

Content Index (C) -- the proportion of content inputs made. This index was calculated for the meeting as a whole. The formula was: $\frac{\Sigma 10 + \Sigma 20 + \Sigma 50 + \Sigma 60}{\Sigma 1 + \Sigma 2 + \Sigma 3}$.

PROCESS INDICES:

Functional Process Index (FP) -- the proportion of functional process inputs made, excluding reinforcement. The FP index was computed for each speaker and for the meeting as a whole by summing the frequencies for categories 72 + 79 and dividing by the total number of inputs. For example, the formula for the resource person was: $\frac{\Sigma 172 + \Sigma 179}{\Sigma 1}$.

Dysfunctional Process Index (DfP) -- the proportion of dysfunctional process inputs made. The DfP index was computed for each speaker and for the meeting as a whole by summing the frequencies for categories 82 + 86 and dividing by the total number of inputs. For example, the formula for the resource person was: $\frac{\Sigma 182 + \Sigma 186}{\Sigma 1}$.

MODE INDICES:

Indirect Structuring Index (I) -- the proportion of inputs made which influenced the discussion indirectly by asking questions and reinforcing other people. The I index was calculated for each speaker by summing the frequencies for categories 10 + 50 + 71 and dividing by the total number of inputs. For example, the formula for the resource person was: $\frac{\Sigma 110 + \Sigma 150 + \Sigma 171}{\Sigma 1}$.

Direct Structuring Index (D) -- the proportion of inputs made which directly influenced the discussion by providing content inputs. The D index was calculated for each speaker by summing the frequencies for categories 20 + 60 and dividing by the total number of inputs. For example, the formula for the resource person was: $\frac{\Sigma 120 + \Sigma 160}{\Sigma 1}$.

GENERAL INDICES:

Activity Level (Act.) -- the proportion of inputs made by each speaker. This index was calculated for each speaker. For example, the formula for the resource person was: $\frac{\Sigma 1}{\Sigma 1 + \Sigma 2 + \Sigma 3}$.

Group Conversation (Grp.) -- the proportion of inputs made by more than one person simultaneously. The formula was: $\frac{\Sigma 4}{\Sigma 1 + \Sigma 2 + \Sigma 3 + \Sigma 4}$.

Sidetracking Index (S) -- the proportion of miscellaneous inputs made during a meeting. The formula was: $\frac{\Sigma 98}{\Sigma 1 + \Sigma 2 + \Sigma 3}$.

Other Instruments

Prior to the start of the survey feedback meetings, two questionnaires were administered. The organizational functioning of the work groups was assessed using a machine-scored, standardized instrument entitled the Survey of Organizations (SOO). A description of the complete instrument together with statistical information regarding the validity and reliability of its component items is provided by Taylor & Bowers (1972).

The questionnaire includes 16 indices and measures five major aspects of organizational life--Climate (pervasive conditions), Supervisory Leadership (the behaviors of a supervisor toward his/her subordinates), Peer Leadership (the behaviors of subordinates toward one another), Group Process (how group members work as a group), and Satisfaction. Most of the questions ask the respondent to rate the extent to which behaviors, conditions, and practices in each of the five areas were characteristic of their work setting (see Appendix D).

The SOO was used in a variety of ways in this study. SOO data were collected at the outset of the OD effort and then fed back to work groups for discussion in the feedback meetings. Approximately 18 months after the start of the OD project, the SOO was re-administered. These two waves of SOO data provided a "pre" and "post" measure of functioning for a subset of groups whose composition remained relatively constant over the 18 months (N = 11 groups). These pre and post measures were used to classify the groups as "successful" or "unsuccessful." Included in the latter category were those groups that did not change and those that changed for the worse. The procedure for classifying each group included five steps:

1. Scores on each of the 16 indices for each group was judged as increasing, remaining the same, or decreasing based on the direction of change from the first to the second survey.

2. For each group, a count was made of the number of indices for which the scores increased, remained the same, or decreased, and the predominant direction of change across indices was noted.

3. Significance of the number of indices moving in the predominant direction was then evaluated using the Sign Test (Siegel, 1956, pp. 68-75).

4. Where the confidence of change in either direction was beyond the five percent level, the group was judged to have moved in that direction ("increased" or "decreased"). Where the significance of the predominant direction was not established by this criterion, the group was judged as not changing.

5. The five groups where the sign tests were significant in the direction of an increase were placed into the "successful" category. The six remaining groups were classified as "unsuccessful."

Comparisons were made between these two categories of groups across the characteristics measured by the coding schemes.

Finally, the SOO measures were viewed as one set of ratings--by groups of subordinates--of various work setting characteristics that could be compared to ratings of similar characteristics by other raters (e.g., the coders, the supervisors). An instrument, entitled the Supervisory Self-Rating Form, was used in comparing supervisors' ratings with other ratings of leadership behaviors. This instrument included questions paralleling the Supervisory Leadership items in the SOO. Supervisors were asked to answer each of the questions as they pertained to their own supervisory behaviors. The supervisors filled in these forms prior to receiving the SOO results for their group feedback session.

Preliminary Findings

Three basic issues (in addition to the coder reliability on the BCS and SRS) were explored for this report:*

1. the range and variability of behaviors measured by the BCS and SRS,
2. the strength of relationships among different measures of similar behaviors, and
3. the existence of relationships between behaviors during the early meetings and "outcomes."

The analyses and results pertaining to each issue are presented below. The smallness of the present sample--smallness in terms of both number of groups and number of meetings--limits the generalizability and conclusiveness of the findings. The results, therefore, may be viewed as suggestive, but preliminary.

Range and Variability of Behaviors

If the BCS and SRS are to be used to identify feedback meeting characteristics which affect the success of survey feedback, the schemes must be sensitive to variations--as well as similarities--in behaviors and processes. Table 6 presents descriptive statistics on the BCS and SRS measures. The data are encouraging on at least three points:

- There is a substantial range in the mean scores for the measures. The means on the BCS indices range from .006 to .600. The means on the SRS indices and single items range from 1.16 to 3.73 for the first meeting and from 1.22 to 4.44 for the second meeting.

---

*Because so few (N = 4) second meetings were coded using the BCS, only the SRS measures were used in analyses for second meetings. Both BCS and SRS measures were used for first meetings.

Table 6

DESCRIPTIVE STATISTICS ON BCS AND SRS MEASURES

	N	MEAN	STD. DEV.	MINIMUM	MAXIMUM
BCS MEASURES (Mtg. 1)					
RP (PI)	16	.54	.17	.09	.91
Sup (PI)	16	.49	.10	.33	.67
Grp (PI)	16	.12	.07	.02	.27
RP (PS)	16	.11	.09	0	.39
Sup (PS)	16	.06	.06	0	.17
Grp (PS)	16	.06	.06	0	.22
Mtg. (PS/C)	16	.11	.09	.004	.36
Mtg. (C)	16	.60	.06	.50	.68
RP (FP)	16	.14	.07	.07	.35
Sup. (FP)	16	.14	.05	.07	.21
Grp. (FP)	16	.04	.03	.01	.14
Mtg. (FP)	16	.10	.03	.05	.16
RP (DfP)	16	.01	.01	0	.05
Sup. (DfP)	16	.02	.02	0	.08
Grp. (DfP)	16	.02	.02	0	.08
Mtg. (DfP)	16	.06	.03	.02	.12
RP (I)	16	.36	.10	.16	.52
Sup. (I)	16	.37	.09	.21	.50
Grp. (I)	16	.29	.09	.19	.46
RP (D)	16	.42	.14	.26	.73
Sup. (D)	16	.39	.12	.23	.59
Grp. (D)	16	.55	.09	.39	.67
RP (Act.)	16	.19	.12	.04	.36
Sup. (Act.)	16	.34	.11	.16	.60
Grp. (Act.)	16	.46	.09	.33	.67
Mtg. (Grp)	16	.08	.04	.03	.14
Mtg. (S)	16	.09	.04	.03	.16
SRS MEASURES (Mtg. 1)					
SRS 1: RP/Makes Impact	19	3.18	1.04	1.00	5.00
SRS 2: RP/Over Active/ Directive	19	2.96	1.24	1.00	5.00
SRS 3: RP/Confronts	19	1.42	.77	1.00	4.00
SRS44: RP/Opposed	19	1.51	.61	1.00	3.00
SRS 5: Sup. Leadership	19	3.03	.57	1.67	4.33
SRS 6: Sup. Prevents Discussion	19	2.05	.88	1.00	4.50
SRS 7: Grp. Partici.	19	3.10	.70	2.00	4.75
SRS 8: Grp. Not Defensive	19	2.60	.57	1.50	3.50
SRS 9: Productive Mtg.	19	3.00	.64	2.00	4.33

Table 6 Continued

	N	MEAN	STD. DEV.	MINIMUM	MAXIMUM
SRS 10: Meeting Unsyst.	19	2.37	.66	1.50	3.50
SRS 11: Objectives Opp.	19	1.25	.97	0	3.00
SRS 12: Problems Outside Control	19	3.37	.73	1.75	4.75
SRS 13: Grp. Not Listening	19	2.24	.65	1.00	3.50
Q3: Grp. Requests Inputs	19	1.57	.69	1.00	3.00
Q5: RP/Info. Resource	19	3.32	1.45	1.00	5.00
Q8: RP/Clear	19	3.74	.81	2.00	5.00
Q12: RP/Intervenes in Emot. Encounters	19	1.16	.37	1.00	2.00
Q16: Double Messages	19	1.53	.70	1.00	3.00
Q49: Grp/Few Dominate	19	2.47	1.02	1.00	4.00
Q22a: Affect Problems	19	3.05	1.22	1.00	5.00
Q22b: Info. Problems	19	2.89	1.20	1.00	5.00
Q22c: Skill Problems	19	1.63	.83	1.00	4.00
Q25: Commit. to Solutions	19	1.74	1.52	0	4.00
Q29: Frustration	19	2.16	.69	1.00	4.00
Q41: Silence	19	1.47	.70	1.00	3.00
SRS MEASURES (Mtg. 2)					
SRS 1	18	3.14	.80	2.00	4.50
SRS 2	18	3.08	1.00	1.25	4.50
SRS 3	18	1.53	.50	1.00	2.50
SRS 4	18	1.42	.34	1.00	2.33
SRS 5	18	3.01	.84	1.33	4.33
SRS 6	18	2.22	.97	1.00	4.50
SRS 7	18	3.08	.92	1.75	4.75
SRS 8	18	2.83	1.00	1.50	5.00
SRS 9	18	2.86	.88	1.33	4.17
SRS 10	18	2.89	.95	1.00	4.50
SRS 11	18	1.22	1.18	0	3.00
SRS 12	18	3.29	.73	2.50	5.00
SRS 13	18	2.44	1.06	1.00	4.50
Q3	18	1.44	.62	1.00	3.00
Q5	18	3.17	1.10	2.00	5.00
Q8	18	4.00	.77	3.00	5.00
Q12	18	1.72	.83	1.00	3.00
Q16	18	1.83	.71	1.00	3.00
Q49	18	2.11	1.08	1.00	4.00
Q22a	18	3.44	1.20	1.00	5.00
Q22b	18	2.44	.98	1.00	4.00
Q22c	18	1.50	.86	1.00	4.00
Q25	18	1.56	1.46	0	4.00
Q29	18	2.33	1.03	1.00	4.00
Q41	18	1.72	1.02	1.00	4.00

- There is a substantial range in the standard deviations (s.d.) for the measures. The s.d.'s on the BCS indices range from .01 to .14. The s.d.'s on the SRS measures range from .37 to 2.57 for the first meeting, and from .34 to 2.15 for the second meeting.

- In many cases the minimum and maximum scores cover a wide range of the possible responses. On 13 out of the 25 SRS measures (i.e., 52%) the differences between the minimum and maximum scores is at least 3.00. (In most cases, the maximum difference possible is 4.00). This holds for both first and second meetings.

Thus, the data indicate that the BCS and SRS measured a range of characteristics, some of which were relatively similar across groups and many which varied across groups.

A similar question concerns whether the SRS and BCS measures were sensitive to similarities and differences across people. To explore this issue, one-way analyses of variance were performed on measures of resource person (RP) behavior for the five internal resource people. The results are shown in Table 7.

As indicated in the table, the five RP's differed significantly ($p<.05$) on four out of 15, or about one-fourth, of the measures. There were significant differences in RP activity levels (two measures) and the extent to which the RP's were clear and acted as information resources. During the second meetings, there were, once again, significant differences ($p<.05$) on two out of eight, or one-fourth, of the measures. During these meetings, the RP's differed in the extent to which they were overactive and the extent to which they confronted group members. Thus, more similarities than differences in RP behavior were tapped by the coding schemes. Some variations were identified, however.

Table 7

ANALYSES OF VARIANCE OF RP BEHAVIORS BY RP

	df	F-Ratio	Level Of Significance
BCS MEASURES (Mtg. 1)			
RP (PI)	4, 11	1.60	p<.24
RP (PS)	4, 11	2.30	p<.12
RP (FP)	4, 11	.68	p<.62
RP (DfP)	4, 11	.91	p<.49
RP (I)	4, 11	.73	p<.60
RP (D)	4, 11	1.77	p<.20
RP (Act.)	4, 11	12.16	p<.001
SRS MEASURES (Mtg. 1)			
SRS 1: RP/Makes Impact	4, 14	1.39	p<.29
SRS 2: RP 'Over Active/ Directive	4, 14	13.41	p<.001
SRS 3: RP/Confronts	4, 14	.33	p<.85
SRS 4: RP/Opposed	4, 14	.49	p<.75
Q3: Grp. Requests Inputs	4, 14	.66	p<.63
Q5: RP/Info. Resource	4, 14	12.78	p<.001
Q8: RP/Clear	4, 14	8.43	p<.001
Q12: RP/Intervenes in Emotional Encounters	4, 14	.41	p<.80
SRS MEASURES (Mtg. 2)			
SRS 1	3, 14	.62	p<.61
SRS 2	3, 14	3.51	p<.04
SRS 3	3, 14	4.92	p<.01
SRS 4	3, 14	1.22	p<.34
Q3	3, 14	.06	p<.98
Q5	3, 14	1.27	p<.32
Q8	3, 14	2.08	p<.15
Q12	3, 14	1.41	p<.28

Relationships Among Multiple Ratings of Behaviors

The data in this study were limited in terms of the number of cases, yet rich in terms of the number and type of variables measured and the measurement procedures used. Two somewhat different ratings of behaviors in the same situation were provided by the BCS and the SRS. In addition to coders' ratings of behavior during meetings, however, the SOO provided subordinates' ratings and the Supervisory Self-Rating Form provided supervisors' ratings of general functioning in the work situation. In the following pages selected comparisons among these ratings are made.

Interrelationships of the Coding Schemes. The first set of comparisons were made in order to explore the interrelationships of the SRS and BCS. The coding schemes provide two slightly different ratings of behavior in the same situation. The number of "comparable" behaviors measured were somewhat limited. Nevertheless, a few comparable measures were available for each of four aspects of the meetings--i.e., the character of the meeting as a whole, and behaviors of the resource person, the supervisor, and the group members.

Table 8 lists the sets of measures compared and shows the correlations between each set. Overall, the correlations were moderately high--ranging from |.10| to |.811| for characteristics of the meetings as a whole, the resource person, and the group members. The median correlation was .434. The measures of supervisory behaviors, however, were only weakly related, with a median correlation of -.082. The correlations of supervisory behaviors ranged from |.018| to |.281|.

Table 8

CORRELATIONS BETWEEN SELECTED BCS AND SRS MEASURES

Measures	Correlation (N = 16)
MEETING CHARACTERISTICS	
SRS 9: Productive Mtg. & Mtg (PS/C)	.492*
& Mtg. (C)	.439
SRS 10: Discussion Unsystematic & Mtg. (S)	.263
RESOURCE PERSON CHARACTERISTICS	
SRS 2: RP/Over Active/Directive & RP (Act.)	.811**
Q5: RP/Info. Resource & RP (I)	.522*
& RP (D)	-.458
SRS 1: RP/Makes Impact & RP (I)	.379
& RP (D)	-.473
Q53: RP/Directs Discussion & RP (I)	.223
& RP (D)	-.435
SUPERVISOR CHARACTERISTICS	
Q48: Sup. Dominates & Sup (Act.)	-.082
Q43: Sup. Open & Sup. (FP)	.075
& Sup.)DfP)	.281
Q44: Sup. Team Builds & Sup. (FP)	-.106
& Sup. (DfP)	.188
Q51: Sup. Directs & Sup (I)	.018
& Sup. (D)	.032
GROUP MEMBER CHARACTERISTICS	
Q42: Battle for Airtime & Mtg. (Grp.)	.559*
Q38: Grp. Participation & Grp. (Act.)	.434
Q37: High Quality Interaction & Grp. (FP)	-.100
& Grp. (DfP)	-.111
Q52: Grp. Directs & Grp (I)	-.147
& Grp (D)	.210

* = $p<.05$

** = $p<.01$

In addition to the reasonable strength of many of the relationships, the correlations were usually in the appropriate direction. For example, the SRS measure "Productive Meeting" is positively related to the BCS measures of problem-solving and content inputs made during the meeting ($v = .492$ and $.439$ respectively). The measures of "activity" have especially high intercorrelations. The BCS measure of the RP's activity level is correlated .811 ($p<.05$) with the SRS measure "RP Overly Active/ Directive." The BCS measure of group conversation is correlated .559 ($p<.05$) with the SRS measure "Battling for Air Time." And the BCS measure of group member activity level is correlated .434 with the SRS measure "Group Participation."

There were also some interesting correlations between the BCS and SRS measures of the resource persons' behaviors. Direct Structuring was defined in the BCS as providing problem-identification or problem-solving inputs while indirect structuring was defined as seeking these inputs and reinforcing other people's inputs. The correlations between the two coding schemes indicate, however, that the BCS measure of direct structuring by the RP is negatively related to the SRS measures RP/information resource ($r = -.458$), RP/directs discussion ($r = -.435$), and RP/makes impact ($r = -.473$). The BCS measure of indirect structuring, on the other hand, is positively related to these three SRS measures. Thus, the RP appears to have made more impact and to have "directed" the discussion more by asking questions than by answering them. This indicates that "Providing Content Inputs" might be a more descriptive index name than the present "Direct Structuring" label for the BCS measure.

There is a range, then, in the strength and direction of the relationships between SRS and BCS measures. A more precise test of how these two types of coding schemes are interrelated would be possible, however, by adding some measures to the SRS which more closely parallel the BCS measures. The stronger correlations were found between the most parallel or comparable questions (e.g., activity levels).

Multi-Method Ratings of Supervisory Leadership. As stated previously, three ratings of supervisory leadership were gathered--the SRS ratings by the coders, the SOO ratings by the subordinates, and the self-ratings by the supervisors. Four major dimensions of supervisory leadership were tapped by all three sets of ratings:

Support--the extent to which the supervisor lets subordinates know that they are worthwhile persons doing useful work.

Team Building--the extent to which the supervisor's behavior encourages subordinates to develop close, cooperative working relationships with one another.

Goal Emphasis--the extent to which the supervisor's behavior stimulates a contagious enthusiasm among his subordinates for doing a good job (not pressure).

Work Facilitation--the extent to which a supervisor's behavior removes roadblocks to doing a good job (Bowers & Seashore, 1966).

The questions used to measure each dimension are listed in Table 9. SOO and self-rating measures were indices comprised of two or more items. The SRS measures were single items.

Table 10 shows the correlations among these three sets of ratings. While the N's were very small and the correlations were not statistically significant for the most part, some interesting patterns emerge. First, the coders' ratings (for both meetings) and the SOO measures were negatively

Table 9

SUPERVISORY LEADERSHIP MEASURES

	SRS MEASURES	SOO MEASURES	SUP. SELF-RATINGS
SUPPORT	To what extent did the supervisor seem open to the opinions and ideas by his/her subordinates?	How friendly and easy to approach is your supervisor?	How friendly and easy to approach are you in you work setting?
		When you talk to your supervisor, to what extent does he pay attention to what you're saying?	When your subordinates ta to you, to what extent do you pay attention to what they're saying?
		To what extent is your supervisor willing to listen to your problems?	To what extent are you willing to listen to your subordinates probl
TEAM BUILDING	To what extent did the supervisor encourage group members to work together as a team?	To what extent does your supervisor encourage the persons who work for him to work as a team?	To what extent doeyou encourage persons who work for you to work as a team?
		To what extent does your supervisor encourage people who work for him to exchange opinions and ideas?	To what extent do you encourage people who work for you to exchang opinions and ideas?
GOAL EMPHASIS	To what extent did the supervisor emphasize goals (work, change objectives, etc.)? (Not pressure)	How much does your supervisor encourage people to give their best effort?	How much do you encourage your subordinates to give their best effort?
		To what extent does your supervisor maintain high standards of performance?	To what extent do you maintain high standards of performance?

Table 9 Continued

	SRS MEASURES	SOO MEASURES	SUP. SELF-RATINGS
	To what extent did the supervisor help to remove roadblocks to solving problems?	To what extent does your supervisor show you how to improve your performance?	To what extent do you show your subordinates how to improve their performance?
WORK FACILITATION		To what extent does your supervisor provide the help you need so that you can plan and schedule work ahead of time?	To what extent do you provide your subordinate with the help they need so they can schedule work ahead of time?
		To what extent does your supervisor offer new ideas for solving job-related problems?	To what extent do you offe new ideas for solving job-related problems?

Table 10

INTERCORRELATIONS OF SUPERVISORY LEADERSHIP MEASURES

Correlations Between SRS Measures (Mtg. 1 and 2) and

	SOO Measures		Supervisor's Self-Ratings	
	Mtg. 1 (N = 19)	Mtg. 2 (N = 18)	Mtg. 1 (N = 16)	Mtg. 2 (N = 18
Support	.184	.118	-.259	-.358
Team Building	.411	.196	-.052	-.373
Goal Emphasis	.060	.154	-.035	-.377
Work Facilitation	-.048	.121	-.034	-.377

Correlations Between SOO Measures and

	Supervisor's Self-Ratings (N = 23)
Support	.020
Team Building	-.310
Goal Emphasis	-.282
Work Facilitation	-.410*

* = $p<.05$

related to supervisors' self-ratings. The coders' ratings and the SOQ measures were, on the other hand, positively but very weakly related. It would be interesting and fruitful to explore these relationships with a larger sample to test whether (1) "third-party" ratings are, in fact, only weakly related to perceptions of subordinates; whether (2) ratings of average supervisory behavior are only weakly related to supervisory behavior in specific settings--e.g., feedback meetings; and whether (3) supervisors' ratings of themselves are often the reverse of others' ratings of them.

Differences Between Successful and Unsuccessful Groups

The question of which feedback meeting characteristics affect the ultimate success of survey feedback, while of considerable interest, cannot really be investigated with the present data set. A modified question may be posed, however:

> Do the meetings of groups that engage in survey feedback successfully differ from the unsuccessful groups as early as the first two feedback meetings?

This question focuses on identifying characteristics of meetings which are either very fixed or very determining, since their presence at the outset would distinguish between the "successful" and "unsuccessful" groups defined several months later. To answer this question, Student's T-Tests were performed. The mean scores on the BCS and SRS measures for the "successful" versus the "unsuccessful" groups were compared.*

---

*The procedure for classifying groups as successful or unsuccessful was described on page 40.

The results of the t-tests are presented as Appendix C. It seems sufficient to say here that the number of characteristics on which the successful and unsuccessful groups differed significantly could easily have occurred by chance. There were no statistically significant differences ($p<.05$) between the two sets of groups during the first meeting, and only one (out of 25) comparison was significant for the second meeting.* This finding indicates that groups are not doomed to failure or assured of success based on the way their first two feedback meetings progress.

---

*The successful groups had less participation by group members during the second meeting than the unsuccessful groups. This may indicate that the successful groups had more room for improvement at the beginning of the OD effort.

Summary and Suggestions for Further Research

Two coding schemes were developed and subjected to preliminary analyses for this study. The Behavior Classification Scheme (BCS) included eleven behavior and four speaker categories which were used to code each problem-solving input made during a survey feedback meeting. The Summary Rating Scheme (SRS) included several survey questions regarding the feedback meeting which coders answered after listening to an entire meeting.

The analyses performed suggest that the two schemes are feasible tools for documenting what transpires during survey feedback sessions:

- their coding reliability and validity is acceptable;
- the schemes seem sensitive to both similarities and differences in meetings and consultant styles; and
- there are some moderate correlations between comparable SRS and BCS measures.

In addition to their potential use as documentation tools, however, the schemes may be used to gather "third party" perceptions of a variety of organizational variables. In this case, coder ratings of supervisory leadership were used in conjunction with subordinate ratings and supervisors' self-ratings to assess the relationships among multiple ratings of supervisory leadership. The results were somewhat surprising. Supervisors' self-ratings tended to be the reverse of ratings provided by both their subordinates and the coders. Coders' ratings were positively but only weakly related to subordinate ratings. While the sample was quite small,

the pattern of correlations indicate that the three ratings provide quite different information.

Future efforts to study feedback meetings using the BCS and SRS should involve at least two extensions of the present research. One extention focuses on reducing the coding time required for each meeting. In the present study, the audio tapes of entire meetings were coded. It might be possible in the future, however, to code a few ten-minute segments of each meeting without sacrificing the essence of the meeting. This possibility could be explored by experimenting with different time-sampling methods.

Secondly, the sample of meetings included for each group should cover feedback sessions from the early, intermediate, and advanced stages of the survey feedback effort. In the present study only the earliest meetings were included. Understandably, successful and unsuccessful groups (as defined by change in the SOO scores) did not differ, for the most part, in terms of the behaviors occurring in the first and second feedback meetings. Extending the sample to include later meetings would allow one to observe the paths of and differences between successful and unsuccessful groups over time.

APPENDIX A: THE CODING SCHEMES AND CODING GUIDELINES

FEEDBACK SESSION TAPES

Deck 001

Column

1-4 STUDY NUMBER (1712)

5-7 SITE NUMBER (from ID form)

8-12 GROUP NUMBER (from ID form)

13-14 MEETING NUMBER (from ID form)

15-17 DECK NUMBER (001)

18-19 RESOURCE PERSON NUMBER (from ID form)

01. G
02. M
03. R
04. P
05. D

20-21 DATE OF MEETING--MONTH (from ID form)

Code actual month (01-12)

99. NA

22-23 DATE OF MEETING--DAY (from ID form)

Code actual day of the month (01-31)

99. NA

Column

24-25 DATE OF MEETING--YEAR (from ID form)

73. 1973
99. NA

26-28 LENGTH OF MEETING (MIN.)

Code approximate length of meeting in minutes (001-997)

29-30 GROUP SIZE (from ID form)

Code actual number of group members (not including supervisor) who filled out the S00 questionnaire.

31-32 GROUP FUNCTION (from ID form)

Code predominant function performed by group members.

White Collar

01. Production I--production tasks requiring low skill levels
02. Production II--production tasks requiring moderate skill levels
03. Production III--production tasks requiring high skill levels
04. Quality control--non-supervisory checking or product quality
05. Clerical--includes typing, filing, editing, etc.
06. Line supervision (production areas only)
07. Staff supervision (supervisors in personnel, administrative, or technical areas.
08. Sales/Marketing Supervisor
09. White Collar, NA Specific function

Blue Collar

11. Production I--production tasks requiring low skill levels
12. Production II--production tasks requiring moderate skill levels
13. Production III--production tasks requiring high skill levels
14. Quality Control--non-supervisory checking or product quality
19. Blue Collar, NA specific function

99. NA

Column

33 CODER/CHECK-CODER NUMBER

Coder	1.	BD
	2.	RG
	3.	CS
	4.	PP
	5.	
Check coder	6.	BD
	7.	RG
	8.	RS
	9.	PP
	0.	

CODER ANALYSIS OF GROUP MEETING

Use the 5-point extent scale below to answer the following questions (1-22, 25-41) and on Deck 044 (42-53).

Five-Point Extent Scale

1. To a very little extent
2. To a little extent
3. To some extent
4. To a great extent
5. To a very great extent

Column

I THE RESOURCE PERSON

34 Q1. PROCESS REMARKS: To what extent did the resource person make comments which affected the process of the discussion (e.g., making maintenance inputs)?

35 Q2. STATEMENTS ABOUT GROUP MEMBERS RELATING: To what extent did the resource person make statements about how group members were relating to each other?

36 Q3. GROUP WANTS INPUTS: To what extent did the group request inputs from the resource person?

37 Q4. MORE INPUTS THAN WANTED: To what extent did the resource person make more inputs than the group seemed to want?

38 Q5. RP AS INFORMATION RESOURCE: To what extent did the resource person act as an information resource?

39 Q6. DIAGNOSIS/STIMULATION: To what extent was the resource person active in making diagnostic comments and/or stimulating action steps?

[redacted]

Column

40 Q7. CONFRONTATION BY RP: To what extent did the resource person actively confront group members?

41 Q8. CLARITY: To what extent were the resource person's statements easy to understand? (Clarity of meaning; not quality of tape reproduction!)

42 Q9. RP PRIMARILY ANSWERING QUESTIONS: To what extent did the resource person remain silent except when responding to questions directed toward her/him?

43 Q10. CONFLICT WITH GROUP: To what extent did conflict exist between the resource person and others at the meeting?

44 Q11. RESISTANCE TOWARD RP: When the resource person intervened, to what extent was resistance encountered in response to the intervention?

45 Q12. RP INTERVENES IN EMOTIONAL ENCOUNTERS: To what extent did the resource person intervene in emotional encounters between others at the meeting

46 Q13. RP ATTACKED: To what extent was the resource person attacked by the group for suggesting changes?

II THE GROUP ITSELF

47 Q14. MISUNDERSTANDING EACH OTHER: To what extent did group members (including the supervisor) misunderstand each other?

48 Q15. SUPERVISOR COMPETENT: To what extent was the supervisor competent in leading the discussion?

Column

49 Q16. DOUBLE MESSAGES: To what extent were double messages being sent? That is, to what extent were there discrepancies between a participant's verbal message and his emotional message?

50 Q17. HONESTY ABOUT SELF: To what extent were members (including the supervisor) honest and candid about themselves?

51 Q18. I-HAVE-TO-CHANGE ATTITUDE: To what extent did individuals take an "I'm-the-one-who-has-to-change" attitude?

52 Q19. OTHERS-MUST-CHANGE ATTITUDE: To what extent was a "change-the-other-guy" attitude in evidence?

53 Q20. SYSTEMATIC APPROACH: To what extent was the problem-solving sequence (the content inputs) followed systematically?

54 Q21. SIDE-TRACKING: To what extent did the discussion get side-tracked on to inconsequential topics?

Q22. TYPES OF PROBLEMS DISCUSSED: To what extent were each of the following types of problems th focus of the discussion:

55 22a. Problems related to feelings or affect or values.

56 22b. Problems related to the lack, or inaccuracy, of information.

57 22c. Problems related to the lack of needed skills.

58 22d. Problems related to constraints created by the situation.

59 Q23. TYPE DISCUSSED MOST: Which type of problem did the group discuss most?

1. Affect-related
2. Information-related
3. Skill-related
4. Situation-related
5. Combination of above

ADDITIONS TO DECK 001

Col.	
69	Q 30. MAINTENANCE INPUTS: To what extent were maintenance inputs provided by the resource person, supervisor, and group members?
70	Q 31. CONTENT INPUTS: To what extent were content inputs provided?
71	Q 32. PROBLEMS IDENTIFIED: To what extent were real problems identified?
72	Q 33. ATTEMPTS TO SOLVE PROBLEMS: To what extent were there attempts to solve problems which were identified?
73	Q 34. PROBLEMS SOLVED: To what extent were problems actually solved?
74	Q 35. PROBLEMS WITHIN CONTROL: To what extent were the problems discussed within the control of the work group and supervisor?
75	Q 36. PROBLEMS OUTSIDE CONTROL: To what extent were the problems discussed outside the control of the work group and supervisor?
76	Q 37. QUALITY OF INTERACTION: To what extent was the group interaction of "high quality"?
77	Q 38. PARTICIPATION: To what extent did group members participate in the discussion?
78	Q 39. GROUP INTERACTION: To what extent did group members interact with each other (rather than merely responding to the supervisor)?
79	Q 40. WILLINGNESS TO TALK ABOUT PROBLEMS: To what extent were group members willing to talk about problems?
80	Q 41. SILENCE: To what extent were there long periods of silence during the meeting?

Column

If the group attempted to solve problems during this meeting, answer question 24 and 25; otherwise code 0 in cols. 57-8 and go to question 26.

60 Q24. TYPE ATTEMPTED FIRST: Which type of problem did the group attempt to solve first?

1. Aspect-related
2. Information-related
3. Skill-related
4. Situation-related

0. Inap, group attempted no solution

61 Q25. CONSENSUS ON SOLUTIONS: To what extent was there consensus commitment to the solutions advanced?

Use 5-point extent scale, EXCEPT:

0. Inap, group attempted no solutions

62 Q26. EXPLANATION OF OBJECTIVES: To what extent were change objectives of the program clearly explained to group members?

63 Q27. DISAGREEMENT ON OBJECTIVES: To what extent was there disagreement on change objectives?

64 Q28. RESISTANCE TO OBJECTIVES: To what extent was there resistance to the change objectives?

65 Q29. FRUSTRATION: How much frustration was in evidence?

1. A little frustration
2. Some frustration
3. Considerable frustration
4. A great deal of frustration

Column

66 SEX OF SUPERVISOR (from ID form)

1. Female
2. Male

67 WAS THE ENTIRE MEETING TAPED?

1. Yes
2. No

68 GROUP SIZE - NO. IN ATTENDANCE (from ID form)

DECK 044

Col.	
1-33	ID Information (same as for Deck 001).
34	Q 42. BATTLING FOR AIR TIME: To what extent were people "battling" for air-time? i.e., to what extent were people interrupting each other and talking all at once in order to get their opinions voiced?
35	Q 43. SUPERVISOR OPEN: To what extent did the supervisor seem open to the opinions and ideas of his/her subordinates?
36	Q 44. SUPERVISOR A TEAM BUILDER: To what extent did the supervisor encourage group members to work together as a team?
37	Q 45. SUPERVISOR A FACILITATOR: To what extent did the supervisor help to remove roadblocks to solving problems?
38	Q 46. SUPERVISOR EMPHASIZES GOALS: To what extent did the supervisor emphasize goals (work, change objectives, etc.)? (not pressure)
39	Q 47. SUPERVISOR SUPPRESSES: To what extent did the supervisor suppress group discussion through his/her attitude or actions?
40	Q 48. SUPERVISOR DOMINATES: To what extent did the supervisor dominate the discussion?
41	Q 49. GROUP MEMBERS DOMINATE: To what extent did a few group members dominate the discussion?
42	Q 50. RESOURCE PERSON DOMINATES: To what extent did the resource person dominate the discussion?
43	Q 51. SUPERVISOR DIRECTS: To what extent did the supervisor actively direct the discussion?
44	Q 52. GROUP MEMBERS DIRECT: To what extent did the group members actively direct the discussion?
45	Q 53. RESOURCE PERSON DIRECTS: To what extent did the resource person actively direct the discussion?

FEEDBACK SESSION TAPES

Decks 002-XXX

Column

1-4 STUDY NUMBER (1712)

5-9 GROUP NUMBER

10-11 MEETING NUMBER

12-14 DECK NUMBER--number decks sequentially from 002 to as many as necessary to complete the coding of one meeting.

Starting in columns 18-19, code in order the problem-solving units occurring in the meeting. These are coded in three-digit groups, the first digit indicating the speaker, the second two indicating which type of input the speaker made. In general every change of speaker should generate a new three-digit group, although one speake may generate several consecutive groups. Guidelines and detailed definitions of the input categories are given on the following pages.

When Deck 002 is filled (through col. 80), continue in Deck 003 on the same code sheet, skipping a line between decks to allow room for inserts. If more than 10 decks are needed, continue on a second code sheet (and on more as needed). If there is any space left on the last deck when the meeting is finished, code 0 in the remaining columns of that deck, through col. 80.

Tape Log. For each tape, keep a running log of the meeting on a separate piece of paper, labeled with the site, group, and meeting numbers of the tape. The log will take the place of "other" cards, and is also to provide an "index" to the coded record, so that the coding and the tape can be matched up if necessary (e.g., for check-coding). Each tape will be check-coded on the segment between two entries (usually not consecutive) in the tape log. The log should contain at least three entries from each deck (from 002 on), and an entry for each problem-solving unit tha you are not sure how to code. Each entry should contain the card:column location of the first digit of the coded response and enough of the speaker's words to identify the response on the tape. Indicate whether the response was logged because of codin difficulties. If the meeting runs onto the second side of the tape, enter the card:column location of the first response from the second side.

Column

15 SPEAKER

1. Resource person
2. Supervisor
3. Group member
4. General response

9. Unidentifiable speaker

17-19 TYPE OF INPUT

	REQUEST	PROVISION
CONTENT INPUTS	10. Exposition	20. Exposition
	15. Progress reporting	25. Progress reporting
	30. Problem definition and clarification 31. Problem in whose control	40. Problem definition and clarification 41. Problem within group's control 42. Problem outside group's control 43. Task-related problem 44. Interpersonal conflict problem
	35. Setting priorities	45. Setting priorities
	50. Identifying solutions 51. Evaluating solutions 52. Active choice of solutions 53. Implementing solutions	60. Identifying solutions 61. Evaluating solutions 62. Active choice of solutions 63. Implementing solutions
	FUNCTIONAL	DYSFUNCTIONAL
MAINTENANCE INPUTS	70. Encouragement of participation 71. Reinforcement/acceptance 72. Positive perception 73. Agreement 74. Disagreement 75. Separation 76. Consensus testing 77. Conflict resolution 78. Summarizing 79. Process comments/checking accuracy	82. Negative perception 83. Negative response 84. Non-productive response 85. Non-response to leads 86. Fragmented discussion 87. Suppression
MISC. INPUTS	95. Seeking approval 96. Evidence of misunderstanding 97. Laughter 98. Miscellaneous 99. Unintelligible/NA 00.	

Guidelines for Input Categories

(see also detailed definitions on following pages)

1. In case of strong doubts regarding statements which could be categorized into two different categories, use the following preference scheme:
   a. Exposition (highest priority)
   b. Problem definition and clarification
   c. Maintenance and miscellaneous inputs except #72, 86, 98, and 99
   d. Setting priorities
   e. Identifying solutions
   f. Evaluating solutions
   g. Active choice of solutions
   h. Implementing solutions
   i. Progress reporting
   j. Maintenance inputs #72, 86
   k. Miscellaneous inputs #98, 99 (lowest pirority)
2. When coding a statement about which you have doubts, keep a log with the following information in it:
   a. The group number and the group meeting
   b. The exact wording of the statement
   c. The code number you assigned
   d. The line and column number the code fell in
3. Clarifying or background information frequently accompanies a request for a content input. If it is impossible to understand the request without including the background information, then code the discourse as one unit--i.e., request for a particular content input. If this request can be understood without the background information, then code the discourse as two units--i.e., provision of content input and request for content input.
4. Remember to use categories 41-44 when they are appropriate. They may be used instead of category 40 when this is merited.
5. Exposition emphasizes content. Maintenance stresses procedure.
6. When the speaker is providing new information, do not categorize the statement as summarizing.
7. Categorize rhetorical questions (i.e., the speaker does not expect a response) as the provision of an input.
8. If in the middle of an input, the speaker calls on another individual, code the main input only once and code "the calling on the other individual" as a separate input at the end of the main input.
9. If in the middle of a content input the speaker interrupts himself to provide maintenance inputs other than just calling on another individual, (1) code the input occurring before the interruption, (2) code the interruption, and (3) code the input occurring after the interruption.
10. If an input is interrupted by another speaker and then completed, code the input only once.
11. Partial comments, interrupted thoughts or confusion caused by many people talking should be categorized as "Fragmented Discussion," (86).
12. If discourse is fragmented but it is clear from the context which input occurred, do not code the input as fragmented discussion; instead, code the discourse in the appropriate input category. The coders should be reasonably certain from the context that the code is correct.

INPUT CATEGORIES-DETAILED DEFINITIONS

I. UNIT OF MEASUREMENT

A. PROBLEM-SOLVING INPUT: The primary unit of measurement is a problem-solving input provided during survey data feedback in a work group meeting. This unit is based upon a single and complete content, maintenance, or miscellaneous input as defined in the categories below, regardless or time required to provide the input. Every time a transition to a new speaker, a new unit is noted.

SPEAKER: Whenever there is a shift in speakers, a new unit is noted. There are five notations for speakers:

1 = resource person speaking
2 = supervisor speaking
3 = work group member speaking
4 = general response
9 = unidentifiable person speaking

II. INPUT CATEGORIES

Eight categories (10,15,30,35,50,51,52,53) indicate that the speaker is requesting that a particular content input be provided. Eight parallel categories (20,25,40,45,60,61,62,63) indicate that the speaker is actually providing the particular content input. Categories 70-79 and 83-87 are maintenance categories, with categories 70-79 representing functional inputs and categories 83-87 representing dysfunctional inputs. Categories 95-99 are miscellaneous categories.

A. REQUEST FOR CONTENT INPUTS

10. EXPOSITION: The speaker requests statements which provide general information about or reactions to (a) the development effort, survey methodology, or meaning of the indices, (b) the purpose of the data feedback session(s), or (c) the score on an index or item.

Exs.: "How many companies are included in the national norm file?"

"What are we supposed to do during this feedback session?"

"How did we score on the Human Resources Primacy index?"

15. PROGRESS REPORTING: The speaker requests statements which provide information about (a) which action steps have been taken to solve a problem or (b) whether the action steps taken have solved the problem they were intended to solve.

Exs.: "What has been done so far to get that new form approved?"

"Has the new work assignment sheet helped to clarify what tasks you are expected to complete each week?"

30. PROBLEM DEFINITION AND CLARIFICATION: The speaker requests statements which (a) identify specific, concrete problems indicated by a particular index or item score, (b) clarify a previous statement, or (c) provide a concrete example of a general problem.

Exs.: "What specific problems with our work methods were you thinking of when you rated this item low?"

"When you said 'that old form,' were you referring to the form for increasing coverage for a present policy-holder?"

"Can you give me a specific example of when favoritism was shown?"

31. Problem Within Whose Control: The speaker requests statements which identify the problem as one which that particular work group can solve or as one which another group or department must solve.

Ex.: "Can we change the form ourselves without going to the department head?"

35. SETTING PRIORITIES: The speaker requests statements which indicate the order in which the problems defined should be solved.

Exs.: "Which problems are most important for us to solve and which are least important?"

"Which problem should we try to solve first?"

50. IDENTIFYING POSSIBLE SOLUTIONS: The speaker requests statements which identify possible solutions to a problem in terms of the concrete steps which must be taken.

Ex.: "What would we have to change in order to enlarge your job so that you will have more responsibility?"

51. EVALUATING PROPOSED SOLUTIONS: The speaker requests statements which evaluate the feasibility, attractiveness, or utility of the suggestions for solving a problem.

Exs.: "What are the pros and cons of this suggestion?"

"Which of these proposed solutions do you think is the best one?"

52. ACTIVE CHOICE OF SOLUTION: The speaker requests statements which indicate which of the proposed solutions will be implemented.

Exs.: "Which solution should we go with?"

"Shall we try Joe's suggestion then?"

53. IMPLEMENTING THE SOLUTION: The speaker requests statements which identify a particular person as responsible for taking specific action steps.

Exs.: "What exactly needs to be done to implement this solution?"

"Who will take the responsibility for making sure these things get done?"

B. PROVISION OF CONTENT INPUTS

20. EXPOSITION: The speaker makes statements which provide general information about or reactions to (a) the development effort, survey methodology, or the meaning of indices, (b) the purpose of the data feedback session(s), or (c) the score on an index or item.

Exs.: "Over 24,000 people have completed this same survey."

"The purpose of our meetings is to find out the survey results, to identify problems indicated by the data; and to solve the problems."

"The score on the Decision-Making Practices index is at the 40th percentile."

25. PROGRESS REPORTING: The speaker makes statements which provide information about (a) which action steps have been taken to solve a problem or (b) whether the action steps taken have solved the problem they were intended to solve.

Exs.: "Since our last meeting, I have developed a draft of a form for dividing our work up, and this week I will be giving all of you copies to react to."

40. PROBLEM DEFINITION AND CLARIFICATION: The speaker makes statements which (a) identify specific, concrete problems indicated by a particular item or index score, (b) clarify a previous statement, or (c) provide a concrete example of a general problem.

Exs.: "I think the low score on the Decision-Making Practices index reflects the fact that we have no say in deciding what days we have off."

"When I said 'old form' I meant the policy termination form."

"I think that when Mary was allowed to go home sick without going to the infirmary first, favoritism was being shown for her."

41. Problem Within Group's Control: The speaker makes statements which identify the problem as one which that particular work group can solve.

Ex.: "I think we can decide ourselves whether or not to establish a little committee to work on rearranging our work area."

42. Problem Outside Group's Control: Speaker makes statements which identify the problem as one which someone outside that particular work group must solve.

Ex.: "Restructuring our jobs is not something that we can do ourselves."

43. Task-Related Problem: The speaker makes statements indicating that the problem is a task-related one as opposed to one centering around interpersonal conflicts.

Ex.: "The problem around here is that there is more work to do than there is time to do it."

44. Problem Centers Around Interpersonal Conflict: The speaker makes statements indicating that the problem centers around interpersonal conflict as opposed to being a very task-related problem.

Ex.: "The thing that bothers me most about trying to establish a liaison with the other unit is that I can't stand the secretary over there."

45. SETTING PRIORITIES: The speaker makes statements which indicate the order in which the problems defined should be solved.

Exs.: "These problems are most important for us to solve and these are least important."

"I think we should try to solve this problem first."

60. IDENTIFYING POSSIBLE SOLUTIONS: The speaker makes statements which identify possible solutions to a problem in terms of the concrete steps which must be taken.

Ex.: "If we want to re-arrange our work area, I think we should select a small committee from among us which will find out what allour preferences are, design a floor plan, and present it to us for suggestions and changes."

61. EVALUATING PROPOSED SOLUTIONS: The speaker makes statements which evaluate the feasibility, attractiveness, or utility of the suggestions for solving a problem.

Ex.: "Using work assignment sheets would help to define what I'm supposed to do each day, but it would also make my day so structured that I'd feel hemmed in."

62. ACTIVE CHOICE OF SOLUTION: The speaker makes statements which indicate which of the proposed solutions will be implemented.

Ex.: "I think we should try that solution to the problem."

63. IMPLEMENTING THE SOLUTION: The speaker makes statements which identify a particular person as responsible for taking specific action steps.

Ex.: "If we're going to change our jobs around, first we need a list of all the various tasks all of us perform."

"I volunteer to work on developing a new form for dividing our work up in a flexible way."

"Why don't you and Mary work on a floor plan for re-arranging our work area."

C. PROVISION OF MAINTENANCE INPUTS

Functional Inputs:

70. ACTIVE ENCOURAGEMENT OF PARTICIPATION: The speaker encourages others to ask questions, offer opinions, and discuss issues by explicitly inviting such participation.

Ex.: "If anyone has any questions about this--or anything else we cover today--please feel free to ask them."

71. REINFORCEMENT/ACCEPTANCE: The speaker makes statements which indicate that the individual should continue his behavior tough not necessarily indicating agreement with content.

Exs.: "Yeah."

"Uh-huh."

"Yeah, yeah, that's good" (code as one unit).

"That's a good idea."

"I understand what you mean."

72. POSITIVE PERCEPTION OF OTHERS: Speaker makes statements which indicate a positive view of some other person(s). Use this category only when the unit is not coded as a content input.

Ex.: "I like you. You're a fun person to be with."

73. EXPRESSING AGREEMENT: The speaker makes statements which clearly indicate agreement with a statement or idea.

Exs.: "I agree with that."

"Yeah, I think you're right."

"I feel the same way as you do about that."

74. EXPRESSING DISAGREEMENT: The speaker makes statements which indicate disagreement with an idea or statement.

Exs.: "I don't agree with that."

"I don't think that's necessarily true."

75. SEPARATING IDEA-SEEKING FROM IDEA-EVALUATION: The speaker makes statements requesting that a problem be defined or suggestions for solving problems be presented without anyone evaluating their feasibility, utility, or attractiveness.

Ex.: "Let's hear everyone's ideas about this before we discuss their relative merits."

76. CONSENSUS-TESTING: The speaker requests statements indicating the degree of agreement with an issue or decision.

Exs.: "Do all of you see the staff shortage as a problem?"

"Is there anyone who does not think we should work on solving this problem first?"

"Does everyone agree that this solution is the best one?"

77. CONFLICT RESOLUTION: The speaker makes statements which encourage two or more people to stop disagreeing with each other so vehemently. The speaker's statements indicate that he/she is seeking the source of the conflict or trying to resolve it, not just trying to cool it out.

Ex.: "Let's slow down for a minute and make sure each of you understands what the other person is saying."

Do not include comments such as "Let's not get excited" and "Calm down, not." Code comments like this as process comments (79).

78. SUMMARIZING: The speaker makes statements reviewing the discussion. The speaker is doing more than paraphrasing another speaker; he/she is also integrating previous discussion.

Ex.: "So what has been said os far is that oftentimes two or more people ask you to do work for them and expect it to be done quickly. However, you don't know whose work to do first and this creates confusion and tension for you."

79. PROCESS COMMENTS/CHECKING ACCURACY: The speaker makes statements which help to guide the discussion, get the group back on the subject, keep the group focussed on the data, or check the accuracy or clarity of some statement (speaker's own or someone else's).

Exs.: "I think we've gotten off the track We were talking about problems related to how we work together as a group."

"Let's stay focussed on problems relevant to item 97 in the Group Process index."

"Let's get started now."

"Is that what you were saying?"

"Am I understanding you?"

"Do you understand what I'm saying?" ("Yes" in answer is coded as reinforcement/acceptance-71).

Disfunctional Inputs:

82. NEGATIVE PERCEPTION OF OTHERS: Speaker makes statements which indicate a negative view of some other person(s). It may take the form of an accusation. Use this category only when the unit is not coded as a content input.

Ex.: "I can't stand the way you always listen to everything that Dale and I talk about in our office." (Code 82 if this statement is not the main statement of a problem. If it is the main statement of a problem, code as 44 (interpersonal conflict problem).

83. NEGATIVE RESPONSE: The speaker indicates an inability or unwillingness to respond to a request or perform a task.

Ex.: "That's not my problem."

84. NON-PRODUCTIVE RESPONSE: The speaker makes irrelevant or disruptive statements.

Exs.: "I'm going to Hawaii for my vacation this year."

"Nothing you've said has made any sense at all."

85. NON-RESPONSE TO LEADS: Period of silence following a question or request, ended by (1) speaker clarifying lead, or (2) same or different speaker going off in a different direction. "Thinking time" of short duration (5 seconds or so) should not be coded as non-response. This category is trying to pick up unwillingness to respond, or inability to respond because of the nature of the question or request. Code speaker as "general response," unless a specific speaker was addressed. Then code the speaker's number who did not respond.

Exs.: "Well, what do you people think we should do about this problem?"(Coded 50)...Silence...Same person, "Well, do you think it's something we should just let go?" (Now, silence coded as 485).

"Don't you think it might help if we improved our training procedures?" (Probably coded as 51) ...Silence... Different person,"I think we should have more paid holidays." (Silence coded as 485. Statement following silence coded as 40.)

"Carol, how can we help you to plan ahead better?" (Carol is the supervisor)....Silence... Supervisor says, "Well, when shall we schedule our next meeting." (Silence coded as 285; Statement following silence coded 279).

86. FRAGMENTED DISCUSSION: A period which cannot be categorized because the statement(s) cannot be understood or cannot be separated.

Exs.: "Ah, well..."

Several people speaking at once about different things.

87. SUPPRESSION: The speaker makes statements which inhibit or squash discussion by indicating that certain topics are not to be discussed because of his/her personal views.

Ex.: "I don't want to talk about that."

"I don't want that talked about here."

If the speaker indicates that a topic should not be discussed at a given time because it would mean diverging from the major topic of conversation, do not code as suppression. Instead, code as process comment.

D. MISCELLANEOUS INPUTS

95. SEEKING APPROVAL: The speaker requests statements which indicate that his behavior is acceptable to others.

Ex.: "Am I wrong to want to be respected?"

96. EVIDENCE OF MISUNDERSTANDING: The speaker calls attention to a misunderstanding. (Checking understanding should be coded as a process comment).

Exs.: "I think you misunderstood me."

"I don't think you understood what she said."

"I misunderstood you."

97. LAUGHTER: Laughter which interrupts the flow of conversation.

98. MISCELLANEOUS: The speaker makes statements which cannot be classified under any of the other categories.

99. UNINTELLIGIBLE/NA: A statement cannot be understood because it is inaudible or drowned out by non-human background noise.

APPENDIX B

A Priori Classification of SRS Items

Used for Establishing Coder Reliability

Category I:	Questions that required mental summing and averaging of observable behaviors.
Included Q:	2, 3, 5, 6, 7, 9, 10, 12, 13, 20, 21, 30, 31, 33, 38, 39, 41, 42, 44, 45, 46, 48 to 53.
Category II:	Questions that required the coders to make judgements about the quality of a behavior or contribution.
Included Q:	1, 8, 11, 14, 15, 22a, 22b, 22c, 22d, 25 to 29, 32, 34, 35 to 37, 43, 47.
Category III:	Questions that required the coder to make inferences about the attitudes or desires of group participants.
Included Q:	4, 16, 17, 18, 19, 40

APPENDIX C

COMPARISONS OF SUCCESSFUL AND UNSUCCESSFUL GROUPS ON BCS AND SRS MEASURES

APPENDIX C

	Mean Of Unsuccessful Groups	Mean Of Successful Groups	Student's t	Level Of Significance
BCS MEASURES (Mtg. 1)	(N=6)	N=5)		
RP (PI)	.55	.63	-1.09	p<.30
Sup. (PI)	.49	.55	-.90	p<.39
Grp. (PI)	.13	.11	.35	p<.73
RP (PS)	.10	.05	1.39	p<.20
Sup. (PS)	.04	.05	-.49	p<.64
Grp. (PS)	.06	.03	1.13	p<.29
Mtg. (PS/C)	.10	.08	.61	p<.55
Mtg. (C)	.59	.62	-.86	p<.41
RP (FP)	.10	.14	-1.78	p<.11
Sup. (FP)	.14	.12	.54	p<.60
Grp. (FP)	.04	.03	.25	p<.80
Mtg. (FP)	.09	.08	.35	p<.74
RP (DfP)	.01	.002	1.29	p<.23
Sup. (DfP)	.03	.01	1.54	p<.16
Grp. (DfP)	.04	.01	1.92	p<.09
Mtg. (DfP)	.06	.04	1.33	p<.22
RP (I)	.39	.34	1.00	p<.34
Sup. (I)	.34	.37	-.63	p<.54
Grp. (I)	.32	.32	-.05	p<.96
RP (D)	.41	.44	-.31	p<.76
Sup. (D)	.43	.40	.41	p<.69
Grp. (D)	.55	.53	.35	p<.73
RP (Act.)	.21	.18	.31	p<. 76
Sup. (Act.)	.35	.36	-.17	p<.87
Grp. (Act.)	.46	.45	-.17	p<.87
Mtg. (Grp.)	.07	.07	-.02	p<.98
Mtg. (S)	.07	.11	-1.62	p<.14
SRS MEASURES (Mtg. 1)	(N=7)	(N=6)		
SRS 1	3.00	3.67	-1.03	p<.32
SRS 2	3.14	3.21	-.08	p<.93
SRS 3	1.43	1.75	-.64	p<.54
SRS 4	1.62	1.78	-.43	p<.68
SRS 5	3.21	3.28	-.24	p<.81
SRS 6	2.00	2.33	-.65	p<.53
SRS 7	2.96	2.96	-.02	p<.99
SRS 8	2.64	2.83	-.59	p<.57
SRS 9	2.83	3.19	-.98	p<.35
SRS 10	2.28	2.58	-.78	p<.45
SRS 11	1.19	1.22	-.05	p<.96
SRS 12	3.29	3.08	.50	p<.63
SRS 13	2.00	2.33	-.86	p<.41

	Mean Of Unsuccessful Groups	Mean Of Successful Groups	Student's t	Level Of Significance
SRS MEASURES (Mtg. 1 Cont.)	(N=7)	(N=6)		
Q3	1.71	1.50	..48	p<.64
Q5	3.86	3.17	.78	p<.45
Q8	3.57	3.67	-.19	p<.85
Q12	1.14	1.17	-.11	p<.91
Q16	1.57	1.67	-.25	p<.80
Q49	2.00	3.00	-1.72	p<.11
Q22a	3.14	3.00	.22	p<.83
Q22b	2.86	3.33	-.75	p<.47
Q22c	1.57	1.50	.24	p<.82
Q25	1.14	2.67	-2.13	p<.06
Q29	2.00	2.17	-.45	p<.66
Q41	1.85	1.00	-.45	p<.66
SRS MEASURES (Mtg. 2)	(N=7)	(N=4)		
SRS 1	3.07	3.00	.13	p<.90
SRS 2	2.89	3.56	-1.02	p<.33
SRS 3	1.43	1.75	-.98	p<.35
SRS 4	1.57	1.50	.32	p<.76
SRS 5	3.35	2.63	1.86	p<.10
SRS 6	2.21	3.00	-1.19	p<.26
SRS 7	3.11	2.31	2.51	p<.03
SRS 8	2.93	2.50	.79	p<.45
SRS 9	2.83	2.42	.82	p<.43
SRS 10	2.71	3.00	-.34	p<.74
SRS 11	1.24	.58	.85	p<.42
SRS 12	3.00	2.94	.25	p<.81
SRS 13	2.36	1.87	.90	p<.39
Q3	1.43	2.00	-1.42	p<.19
Q5	3.29	3.75	-.64	p<.54
Q8	3.71	3.75	-.08	p<.93
Q12	1.43	1.25	.54	p<.60
Q16	1.86	1.75	.40	p<.70
Q49	2.00	2.00	0	1.00
Q22a	3.00	2.75	.33	p<.74
Q22b	2.71	3.00	-.50	p<.63
Q22c	1.57	2.00	-.66	p<.53
Q25	1.71	2.00	-.27	p<.79
Q29	2.00	1.75	.55	p<.60
Q49	1.85	1.50	.54	p<.60

APPENDIX D

Sixteen Indices and Component Items

From The Survey of Organizations

I. Decision Making Practices

A. How are objectives set in this organization?

1. Objectives are announced with no opportunity to raise questions or give comments

2.. Objectives are announced and explained, and an opportunity is then given to ask questions

3. Objectives are drawn up, but are discussed with subordinates and sometimes modified before being issued

4. Specific alternative objectives are drawn up by supervisors, and subordinates are asked to discuss them, and indicate the one they think is best

5. Problems are presented to those persons who are involved, and the objectives felt to be best are then set by the subordinates and the supervisor jointly, by group participation and discussion

B. In this organization to what extent are decisions made at those levels where the most adequate and accurate information is available?

C. When decisions are being made, to what extent are the persons affected asked for their ideas?

D. People at all levels of an organization usually have know-how that could be of use to decision-makers. To what extent is information widely shared in this organization so that those who make decisions have access to all available kr[illegible]

II. Human Resources Primacy

A. To what extent does this organization have a real interest in the welfare and happiness of those who work here?

B. How much does this organization try to improve working conditions?

C. To what extent are work activities sensibly organized in this organization?

III. Technological Readiness

A. To what extent is this organization generally quick to use improved work methods?

B. To what extent are the equipment and resources you have to do your work with adequate, efficient, and well-maintained?

IV. Lower Level Influence

In general, how much say or influence does each of the following groups and people have on what goes on in your department?

A. Lowest-level supervisors (foremen, office supervisors, etc.)

B. Employees (people who have no subordinates)

V. Communication Flow

A. How adequate for your needs is the amount of information you get about what is going on in other departments or shifts?

B. How receptive are those above you to your ideas and suggestions?

C. To what extent are you told what you need to know to do your job in the best possible way?

VI. Motivational Conditions

A. How are differences and disagreements between units or departments handled in this organization?

1. Disagreements are almost always avoided, denied, or suppressed

2. Disagreements are often avoided, denied, or suppressed

3. Sometimes disagreements are accepted and worked through; sometimes they are avoided or suppressed

4. Disagreements are usually accepted as necessary and desirable and worked through

5. Disagreements are almost always accepted as necessary and desirable and are worked through

B. Why do people work hard in this organization?

1. Just to keep their jobs and avoid being chewed out

2. To keep their jobs and make money

3. To keep their jobs, make money, and seek promotions

4. To keep their jobs, make money, seek promotions, and for the satisfaction of a job well done

5. To keep their jobs, make money, seek promotions, do a satisfying job, and because other people in their work group expect it

C. To what extent are there things about working here (people, policies, or conditions) that encourage you to work hard?

VII. Supervisory Support

A. How friendly and easy to approach is your supervisor?

B. When you talk with your supervisor, to what extent does he pay attention to what you're saying?

C. To what extent is your supervisor willing to listen to your problems?

VIII. Supervisory Goal Emphasis

A. How much does your supervisor encourage people to give their best effort?

B. To what extent does your supervisor maintain high standards of performance?

IX. Supervisory Work Facilitation

A. To what extent does your supervisor show you how to improve your performance?

B. To what extent does your supervisor provide the help you need so that you can schedule work ahead of time?

C. To what extent does your supervisor offer new ideas for solving job-related problems?

X. Supervisory Team Building

A. To what extent does your supervisor encourage the persons who work for him to work as a team?

B. To what extent does your supervisor encourage people who work for him to exchange opinions and ideas?

XI. Peer Support

A. How friendly and easy to approach are the persons in your work group?

B. When you talk with persons in your work group, to what extent do they pay attention to what you're saying?

C. To what extent are persons in your work group willing to listen to your problems?

XII. Peer Goal Emphasis

A. How much do persons in your work group encourage each other to give their best effort?

B. To what extent do persons in your work group maintain high standards of performance?

XIII. Peer Work Facilitation

A. To what extent do persons in your work group help you find ways to do a better job?

B. To what extent do persons in your work group provide the help you need so that you can plan, organize, and schedule work ahead of time?

C. To what extent do persons in your work group offer each other new ideas for solving job-related problems?

XIV. Peer Interaction Facilitation

A. How much do persons in your work group encourage each other to work as a team?

B. How much do persons in your work group emphasize a team goal?

C. To what extent do persons in your work group exchange opinions and ideas?

XV. Group Process

A. To what extent does your work group plan together and coordinate its efforts?

B. To what extent does your work group make good decisions and solve problems well?

C. To what extent is information about important events and situations shared within your work group?

D. To what extent do persons in your work group know what their jobs are and know how to do them well?

E. To what extent do you have confidence and trust in the persons in your work group?

F. To what extent is your work group able to respond to unusual work demands placed upon it?

G. To what extent does your work group really want to meet its objectives?

REFERENCES

Alderfer, C.P. & Lodahl, T.M. A quasi experiment on the use of experimental methods in the classroom. Journal of Applied Behavioral Science, 1971, 7, 43-70.

Amidon, E.J. & Hough, J.B. (Eds.) Interactional Analysis: Theory, Research and Application. Reading, Mass.: Addison-Wesley Publishing Co., 1967.

Bales, R.F. Interaction content analysis. In Preliminary Report of the First National Training Laboratory on Group Development, National Educational Association and Research Center for Group Dynamics, M.I.T., 1947.

Bales, R.F. Interaction Process Analysis, A Method for the Study of Small Groups. Reading, Mass.: Addison-Wesley,Publishing Company, Inc., 1950a.

Bales, R.F. A set of categories for the analysis of small group interaction. American Sociological Review, 1950b, 15 (2), 257-263.

Bales, R.F. How people interact in conferences. Scientific American, March 1955, 31-35.

Bales, R.F. Task roles and social roles in problem-solving groups. In Maccoby, E.E., Newcomb, T.M., & Hartley, E.L. (eds.), Readings in Social Psychology. New York: Holt, Rinehart and Winston, Inc., 1958, 443-447.

Bales, R.F. Personality and Interpersonal Behavior. New York: Holt, Rinehart, and Winston, Inc., 1970.

Berg, D.N. The Yale Criteria Study Follow-Up. Interview, Content and Process. Unpublished Paper, 1972.

Bowers, D.G. OD techniques and their results in 23 organizations: The Michigan ICL study. Journal of Applied Behavioral Science, 1973, 9 (1), 21-43.

Bowers, D.G. Perspective in organization development. Working Paper, CRUSK, Institute for Social Research, 1970.

Bowers, D.G. & Hausser, D.L. Group types and intervention effects in organizational development. Technical Report to the Office of Naval Research, 1975.

Bowers, D.G. & Seashore, S.E. Predicitng organizational effectiveness with a four-factor theory of leadership. Administrative Science Quarterly, 1966, 11, 238-263.

Cohen, J. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. Psychological Bulletin, 1969, 70, 213-220.

Franklin, J.L. Characteristics of successful and unsuccessful organizational development. Technical Report to the Office of Naval Research, 1975.

Jenkins, G.P. Jr., Nadler, D.A., Lawler, E.E. III, & Cammann, C. Standardized observations: An approach to measuring the nature of jobs. Journal of Applied Psychology, 1975, 60 (2), 171-181.

Klein, S.M., Kraut, A.I., & Wolfson, A. Employee reactions to attitude survey feedback: A study of the impact of structure and process. Administrative Science Quarterly, 1971, 16, 497-514.

Massialas, B.G., Sprague, N.F., & Sweeney, J.C. Structure and process of inquiry into social issues in secondary schools. Technical Report to the Department of Health, Education, and Welfare, 1970.

Siegel, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill Book Company, 1956.

Taylor, J.C. & Bowers, D.G. Survey of Organizations. Ann Arbor, Michigan: Institute for Social Research, 1972.